Developmental epigenetics in humans: Can maternal nutritional status mediate DNA methylation in their offspring?

Paula Dominguez-Salas, MSc

Thesis submitted in fulfilment of the requirement for the award of the degree of Doctor of Philosophy (PhD)

London School of Hygiene and Tropical Medicine

MRC International Nutrition Group

October 2012

To my grandfather, who deserved a PhD himself,

and the rest of my family…

STATEMENT OF OWN WORK

School's definition of Plagiarism and Cheating is as follows (the full definition is given in the Research Degrees Handbook):

'Plagiarism is the act of presenting the ideas or discoveries of another as one's own. To copy sentences, phrases or even striking expressions without acknowledgement in a manner which may deceive the reader as to the source is plagiarism. Where such copying or close paraphrase has occurred the mere mention of the source in a biography will not be deemed sufficient acknowledgement; in each instance, it must be referred specifically to its source. Verbatim quotations must be directly acknowledged, either in inverted commas or by indenting.' (University of Kent)

Plagiarism may include collusion with another student, or the unacknowledged use of a fellow student's work with or without their knowledge and consent. Similarly, the direct copying by students of their own original writings qualifies as plagiarism if the fact that the work has been or is to be presented elsewhere is not clearly stated.

Cheating is similar to plagiarism, but more serious. Cheating means submitting another student's work, knowledge or ideas, while pretending that they are your own, for formal assessment or evaluation. Supervisors should be consulted if there are any doubts about what is permissible.

Declaration by Candidate

I have read and understood the School's definition of plagiarism and cheating given in the Research Degrees Handbook. I declare that this thesis is my own work, and that I have acknowledged all results and quotations from the published or unpublished work of other people.

I have read and understood the School's definition and policy on the use of third parties (either paid or unpaid) who have contributed to the preparation of this thesis by providing copy editing and, or, proof reading services. I declare that no changes to the intellectual content or substance of this thesis were made as a result of this advice, and, that I have fully acknowledged all such contributions.

Signed:
Date: October 30, 2012
Full name: PAULA DOMINGUEZ-SALAS

ACKNOWLEDGEMENTS

Big thanks to every participant of this study for always welcoming our team with a big smile. Without their selfless help the study would have not been possible. My thanks also to the Wellcome Trust for generously funding this project.

This work would have never taken place either without the invaluable help and input from my colleagues. There are not enough words to thank my two supervisors, Drs Branwen Hennig and Sophie Moore, for teaching me many new skills and for their invaluable, precious and tireless guidance. Thanks for your clever and attentive supervision and for being such good bosses, good colleagues and good friends. My heartfelt thanks to my advisory board, Drs Sharon Cox, who introduced me into the methylation world as an MSc student, and to Dorothea Nitsch, who provided an external view on the project and her very valuable hints for my work. To our head of group, Prof Andrew Prentice, for offering me such an extraordinary opportunity and for all his care; to all my ING colleagues for their good advice and for having made my work not only inspiring but also enjoyable and fun: to Tony for always-available-statistics-without-tears (and birds), I will never be grateful enough; to Kerry, Jenny and Myriam for their helpful advice on lab work and administration and for making special every single day in our shared offices; to Mathilde and Sophie H for always being two steps ahead of me and for their friendly and supportive sharing of their experiences on 'PhD-ing'; to Stefan for being always on the same step and sharing the PhD-headaches. To all my other LSHTM, ING and HNR (with special thanks to Dee) colleagues in London, Cambridge and Keneba for their inputs, their time and their support. To every single member of MRC Keneba staff for their help and warmth, and for making my Gambian life exciting. To Maggie and Henry for their relentless help and good company in Keneba. To our incredibly helpful collaborators, Profs Steve Zeisel, Sheila Innis and Rob Waterland, and their teams, for receiving me in

their labs and sharing their experiences. Working with you all has been more than a pleasure.

Last but not least, I thank my Father, who always loves all my ideas, and my Mother, who usually dislikes them but still supports me. To my Grandmother for the candles she lights for my safe trips and my successes in life. To Dr Kwame, for encouraging me into this adventure and for the drawings. My warmest thanks too to the rest of my family and my close friends in London, Spain and elsewhere: to all the others I cannot name here but who are always there for me and encouraging when things go wrong and who will share my happiness when this thesis is finished.

Thanks! Abarakka! ¡Gracias! Jaaraama!

ABSTRACT

Animal studies show that periconceptional maternal supplementation with nutrients involved in the provision of methyl-groups can alter DNA methylation patterns in the offspring. This thesis examines possible associations between maternal nutrition and offspring DNA methylation patterns in humans in the context of seasonal influences.

Two complementary studies were conducted in rural Gambia. The 'indicator' group enrolled 30 non-pregnant women and followed them monthly for one year, to measure dietary intakes and blood biomarker levels of choline, betaine, folate, methionine and vitamins B2, B6 and B12, as well as plasma homocysteine, S-adenosyl-methionine (SAM), S-adenosyl-homocysteine (SAH), cysteine and dimethylglycine (DMG). The 'main' group enrolled 136 mother-infant pairs, with infants conceived at the peak of the rainy (July to September) or dry (February to April) season. The same panel of blood biomarkers were measured on maternal samples collected early in pregnancy. In the infants, DNA methylation at seven metastable epialleles (MEs) was assessed.

Significant seasonal variation was observed within the indicator group for dietary intakes of choline, betaine, folate and B2 and all metabolic biomarkers. A possible change in dependence between the betaine and folate pathways between seasons was identified. Within the main group, a seasonal variation was observed in both maternal biomarker levels and infant DNA methylation. The rainy season was associated with a higher maternal SAM:SAH ratio, higher concentrations of most methyl-donors (folate, betaine and methionine) and cofactors (B2 and B6) and higher infant DNA methylation. DNA methylation at the MEs studied was associated with B2, cysteine and the SAM:SAH and DMG:betaine ratios, but not with the other biomarkers.

These data suggest that Gambian seasonality influences maternal methyl-group supply, which may be linked to DNA methylation patterns in offspring. Future work is required to confirm this observation, and understand precise mechanisms, so optimal nutrition at critical phases can be determined.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AI:	Adequate intake
ANOVA:	Analysis of variance
BMI:	Body mass index
CCV:	Coefficient of cyclic variation
Cho:	Free choline
CFRI/UBC:	Child and Family Research Institute at the University of British Columbia
CpG:	Cytosine-phospate-guanine
CVD:	Cardiovascular disease
CVi:	Coefficient of inter-assay variation
DALY:	Disability-adjusted life-years
DMR:	Differentially methylated regions
DOB:	Date of birth
DINO:	Diet In Nutrients Out
DMG:	Dimethylglycine
DNA:	Deoxyribonucleic acid
DNMT:	DNA methyltransferase
DOHaD:	Developmental Origins of Health and Disease
DSS:	Demographical surveillance system
EAR:	Estimated Average Requirement
EC:	Ethics Committee
EDTA:	Ethylenediaminetetraacetic acid
EGRAC:	Erythrocyte glutathione reductase activation coefficient
ENID:	Early Nutrition and Immune Development trial
FeFol:	Iron and folate supplement
GEE:	Generalised Estimated Equations
GLS:	Generalised least square
GM:	Geometric mean
GPC:	Glycerophosphocholine
HEP:	Human Epigenome project
HGP:	Human Genome project

HPA: Hypothalamic-pituitary-adrenal

ID: Identification number

IGF: Insuline-like growth factor

LAZ: Length-for-age Z-score

LC/ESI IDMS: Liquid chromatography/electrospray ionization-isotope dilution mass spectrometry

LC-MS-MS: Liquid chromatography-tandem mass spectrometry

LMICs: Low- and middle-income countries

LMP: Last menstrual period

LSHTM: London School of Hygiene and Tropical Medicine

MANOVA: Multivariate analysis of variance

MDEG: Methyl-Donors and EpiGenetics study

ME: Metastable epiallele

MRC: Medical Research Council

MRC HNR : MRC Human Nutrition Research

MTHFR: 5,10-methylenetetrahydrofolate reductase

MUAC: Mid-upper arm circumference

NCD: Non-communicable diseases

NTD: Neural tube deffects

PA: 4-Pyridoxic acid

PAH: Polycyclic aromatic hydrocarbons

PCA: Principal component analysis

Pcho: Phosphocholine

PEMT: Phosphatidylethanolamine-N-methyltransferase

PI: Principal investigator

PL: Pyridoxal

PLP: Pyridoxal-5'-phosphate

PTC: Phosphatidylcholine

RBC: Red blood cells

RCT: Randomised controlled trial

RNA: Ribonucleic acid

SAH: S-adenosylhomocysteine

SAM:	S-adenosylmethionine
SCC:	Scientific Coordinating Committee
SGA:	Small for gestational age
SM:	Sphingomyelin
SNP:	Single nucleotide polymorphism
tHcy:	Total homocysteine
UBC:	University of British Columbia
UCL:	University College of London
UNC-CH:	University of North Carolina-Chapel Hill
WHO:	World Health Organisation
WK:	West Kiang
WLZ:	Weight-for-length Z-score
WT:	Wellcome Trust

PART I: INTRODUCTION

"…epigenetic changes, at least in part, explain the link between a poor start to life and later disease risk. It strengthens the case for all women of reproductive age having greater access to nutritional, education and lifestyle support to improve the health of the next generation, and to reduce the risk of the conditions such as diabetes and heart disease which often follow obesity" (University of Southampton's press release, 18th April 2011; Mark Hanson British Heart Foundation Professor and Director of the University of Southampton's Human Development and Health Unit)

CHAPTER 1. BACKGROUND TO THIS THESIS

In this PhD thesis the background, design and findings of the study conducted by myself in rural Gambia from June 2009 through July 2011 are presented. The thesis also includes a final discussion where potential future developments are illustrated and some conclusions for further research are provided.

1.1. THESIS STRUCTURE

This PhD thesis follows a 'mixed style', where research papers are incorporated into broader thesis Chapters. Four stand-alone documents have been included, all written by myself as first author: a book Chapter on epigenetics which is waiting for editor's revision for publication, has been included as part of the introductory literature review. The three other stand-alone documents are scientific papers: the first (literature review on one-carbon metabolism, in Chapter 2, Section 2.2.) is published; the second (results of one of the PhD studies, in Chapter 4, Section 4.1.) has been submitted and comments from reviewers are currently being addressed; and the third (results of the other PhD study, in Chapter 5, Section 5.1.) is shortly going to be submitted, in peer-reviewed journals. The different manuscripts are focused on the same research topic but are submitted as independent research contributions. Inevitably, this has led to some repetition of concepts and methods. In addition, where critical information was not covered in the research papers but was considered relevant to this thesis, additional details are described separately. Each manuscript is prefaced by the required cover sheet.

Part I contains three introductory Chapters: Chapter 1 is an introduction to the thesis, Chapter 2 an introduction to the research topic (including the review paper published and a book Chapter on epigenetics) and Chapter 3 an introduction to the PhD research studies. Part II contains two Chapters, each detailing one of the two study groups investigated as part of this PhD project. Chapters 4 and 5 comprise one research paper each, as well as additional information and results. Part III contains a final discussion and conclusions for both studies in combination (Chapter 6). It also describes directions for future research, as well as outlining future publications to be written, not included in this thesis.

1.2.CANDIDATE'S INVOLVEMENT

I defined the topic of this thesis and the main hypothesis in close collaboration with my two supervisors, Dr Branwen Hennig and Dr Sophie Moore, as well as Prof Andrew Prentice, Dr Sharon Cox and Dr Tony Fulford. The research forms the continuation of a pilot project that I conducted as MSc student in 'Public Health Nutrition' at the London School of Hygiene and Tropical Medicine (LSHTM) in 2007, on intake of methyl-donors in The Gambia. My contribution involved all stages of the Wellcome Trust (WT) grant proposal preparation and study design including definition of methods, liaison with our collaborators and elaboration of the budget.

Once the grant proposal was approved, I started working at the Medical Research Council (MRC) International Nutrition Group (ING) and became the MDEG (Methyl-Donors and Epigenetics) project principal investigator (PI), based primarily at the MRC Keneba field station in The Gambia. There, I took responsibility for the development of all the preparatory documents (field manual, protocols and ethical applications) and for setting up, implementing and monitoring the project data collection. I was based full-time in The Gambia for two years in order to organise, conduct and supervise the field data collection. Prior to the start of the fieldwork, I was in charge of planning the logistics, designing the database and training the study team (field assistants, lab technicians, nurses and data clerks). During the study, I also organised the field visits and lab work, participating in them when possible and then supervising and monitoring the sample collection by the lab and field teams when I was not available to lead. I organised the shipment of samples to collaborators and monitored data entry. Particularly, a great deal of the dietary data entry was done by myself to pilot the new version of the dietary software, and I also contributed with Darren Cole, its designer and programmer, to the software development. The data cleaning was carried out by myself.

All statistical analyses presented in this thesis fell under my responsibility, receiving expert statistical guidance from Dr Tony Fulford. This thesis was entirely written by me and I am first author on all research papers presented as part of it. The papers I wrote were commented on by the relevant co-authors, whose inputs were incorporated prior to submission for publication.

1.3.COLLABORATING INSTITUTIONS

Very specific laboratory work was needed for each of the areas of this project, for which MRC ING did not have the equipment or the expertise. Therefore, at the point of the initial grant application, appropriate collaborators, who were well-known experts in their field were identified. At the stage of the sample analysis, I visited all the laboratory facilities and I was shown or participated in the techniques used, hands-on. These collaborators and their institutions were the following:

- Prof Steven H. Zeisel (Department of Nutrition, Gillings School of Global Public Health and School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA). His team analysed the food sample content for choline and betaine.
- Prof Sheila Innis (Department of Pediatrics, University of British Columbia, Vancouver, British Columbia, Canada). Her team analysed the maternal plasma biomarkers.
- Prof Rob Waterland (Departments of Pediatrics and Molecular & Human Genetics, Baylor College of Medicine, USDA/ARS Children's Nutrition Research Center, Houston, Texas, US). His team analysed the infant DNA methylation.
- Christine Clewes and Lorna Cox (MRC Human Nutrition Research (HNR), Elsie Widdowson Laboratory, Cambridge, UK). They organised the B2 status assays. Additionally, Darren Cole, Dr Gail Goldberg and Dr Celia H Prynne (MRC-HNR) provided the computer tools for the dietary assessment.

1.4.FUNDING

A grant to support this research was awarded by the Wellcome Trust (WT ref. 086369MA) to Dr Branwen Hennig, with Dr Sophie Moore and Prof Andrew Prentice as co-PIs and myself as named researcher. This research grant covered most costs of the project, including my salary and laboratory analysis of substances under study by collaborators and commercial laboratories. MRC core funding to the International Nutrition Group (MC-A760-5QX00) covered additional fieldwork costs in The Gambia.

1.5.STUDY TIMELINE

The chronology of the fieldwork activities and subsequent analysis is shown in Appendix I.

CHAPTER 2: STUDY BACKGROUND

2.1.INTRODUCTION

Evidence for the intra-uterine environment's influence, in early development, on susceptibility to chronic diseases such as cardiovascular disease (CVD) or diabetes during adulthood (1) is now compelling. Maternal undernutrition is one such influence (2), which may be of particular relevance in low- and middle-income countries (LMICs) . As an example, between five and 20% of African women have a low BMI as a result of chronic undernutrition (3). Additionally, it was estimated that more than 20 million infants worldwide (15.5% of all births) are born with low birth weight, and 95.6% of these in LMICs (4). Around 15% of the infants born in sub-Saharan Africa weigh less than 2,500 g (4). In parallel to this, in LMICs undergoing social, nutritional and economic transition, chronic diseases are becoming critical in mortality rates (5): 41 million people are estimated to die of chronic non-communicable diseases (NCDs) in 2015, out of a total of 64 million deaths, and four out of five of these deaths will be in LMICs (6). In addition to mortality rates, chronic diseases remain amongst the top causes of disability-adjusted life-years (DALYs) lost globally. NCDs have a huge negative impact, particularly in LMICs economies, since NCDs usually affect people still economically productive (7). For example, the global health burden of stroke alone is of 44 million DALY's lost (8) and estimated losses to gross domestic product as the result of coronary disease, cerebrovascular disease, and diabetes in 2005 were $20 million dollars in a country such as Ethiopia (8).

The concept of 'mismatch' relates to the metabolic response to fetal nutritional deficits, designed to cope with environments where nutrition is marginal (and potentially providing short-term advantages), but which may then be detrimental when followed by a transition environment with energy excess and lack of exercise (9). This can manifest with

disease and thus can become a key cause of increasing prevalence chronic diseases in LMICs (10). Adaptations to environmental insults during critical windows of development may permanently alter physiological function, through changes in tissue and organ structure, metabolism or gene expression (11). The mechanisms triggering these permanent changes are known as "metabolic programming" or also, given that nutrient supply variations are thought to be one of the stronger environmental stimulus for these adaptive processes, "nutritional programming" (12). The molecular basis for these mechanisms is still unclear. Several candidate mechanisms have been identified, such as: 1) permanent structural changes in an organ due to suboptimal concentrations of key factors during development (e.g. reduction in ß cells in the endocrine pancreas); 2) persistent epigenetic modifications altering gene expression (e.g. DNA methylation); and 3) permanent effects on regulation of the cellular aging (e.g. increases in oxidative stress) (13), which might interact to a different extent. Disentangling and understanding such mechanisms is paramount to design appropriate prevention interventions.

This PhD is aiming at contributing to this broad goal, by focusing on metabolic programming, specifically DNA methylation, an epigenetic mechanism which has attracted a lot of attention in the recent past due to its role in cancer and other chronic conditions as well as early development (14). DNA methylation consists of the addition of methyl-groups to the DNA chain, a process that interferes with transcription. Methyl-groups utilised in DNA methylation (described in detail in Section 2.2.2.) come ultimately from the diet. This thesis investigates whether the links observed between maternal dietary methyl-donor supply and epigenetic outcomes in the offspring in animal models also operate in man, thus aiming to increase our knowledge of epigenetic processes occurring during early human development.

This Chapter contains a literature review of my research topic (Section 2.2.), with sections from a book Chapter which introduces the different epigenetic mechanisms

(Section 2.2.1) and a published review paper that provides further detail on DNA methylation, the links with nutrition and the existent evidence of DNA methylation in fetal programming (Section 2.2.2.). Section 2.3 includes the background information about the study setting (West Kiang, The Gambia), and Sections 2.4 and 2.5. introduce my PhD research studies.

2.2.LITERATURE REVIEW

2.2.1. BOOK CHAPTER – 'Epigenetics'

Article cover sheet:

1. For a 'research paper' prepared for publication but not yet published

   1.1. Where is the work intended to be published?

   This manuscript is a book Chapter on Epigenetics written for publication in a textbook on "Basic principles of genetic epidemiology in human and infectious diseases". The purpose of this Chapter is to give an introductory overview to different epigenetic mechanisms and their importance in human developmental physiology and disease. As there is substantial overlap with other sections included in this thesis, two full sections have been eliminated from the book Chapter, namely those sections which describe the dependence of methylation on dietary substrates (one-carbon metabolism) and the role of epigenetics in fetal programming, areas thoroughly discussed in the review paper I in Section 2.2.2. Other sections in this Chapter go beyond the scope of this thesis, as they describe in broad detail the epigenetic diseases and therapy, and therefore, I have also excluded them from this thesis. For the easy flowing of the text, the rest of the book Chapter sections have remained untouched, even if some repetition still exists.

   1.2. List the paper's authors in the intended authorship order

   Dominguez-Salas P and Moore G

   1.3. Stage of publication – Not yet submitted/Submitted/**Undergoing revision from editor's comments**/In press

   It was submitted on 01/10/2010.

2. For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

   I was approached by the editors to write this book Chapter. I led the writing phase and it was then reviewed by Prof Gudrun Moore (Institute of Child Health, University College of London, UK). The Chapter is currently awaiting feedback from the editors. It reviews epigenetic mechanisms in a broader sense, with special attention to DNA methylation, and their health and disease consequences.

Candidate's signature _____

Supervisor or senior author's signature to confirm role as stated in (2) _____

EPIGENETICS

Paula Dominguez-Salas[1], Gudrun Moore[2]

[1] MRC International Nutrition Group, London School of Hygiene & Tropical Medicine, London, UK
[2] Institute of Child Health, UCL, London, UK

The 'Human Genome project' (HGP) was launched in 1990 (15) to sequence the genome and to map the approximately 19,600 genes contained within it (16). By 2003 the sequence of 3 million base pairs had been obtained. Now the actual biological functions of these genes need to be elucidated. After this wonderful breakthrough in 2003, scientists launched the next stage in gene regulation discovery, the delineation of the epigenome. This is the study and mapping of the layer of information or chemical modifications on top of the DNA bases, found to be key in both regulation and function of the genes beneath.

In 1950 Waddington (17) defined the term 'epigenetics' as "any mechanism by which multicellular organisms with identical genotype in all cells could develop diverse expression patterns or phenotype". The modern definition describes epigenetics as "the study of stable and heritable alterations in gene expression potential, which is not mediated by DNA sequence alteration" (18, 19). Epigenetic marks can even be inherited through generations (transmitted from parent to offspring via the gamete) and this has been shown in both plants and mice (20). However, apart from the well characterised imprinted genes (discussed in Section 2.2.1.) the transgenerational epigenetic inheritance has not been shown in humans (21).

2.2.1.1. Molecular Mechanisms in Epigenetics

The basis of epigenetics is complex, bringing together different additive molecular layers collectively referred to as epigenome. There are many different molecular modification or mechanisms that fall under the term 'epigenetics' with common agreement in the inclusion of DNA methylation, histone modifications, antisense RNAs and small interfering RNAs. All these different epigenetic mechanisms physically regulate DNA transcription of specific genes for expression or silencing, without altering the primary DNA sequence. Depending on the gene, epigenetic disturbance may lead to different phenotypic outcomes, some physiologically beneficial (i.e. oncogene suppression) and others detrimental (i.e. developmental syndromes).

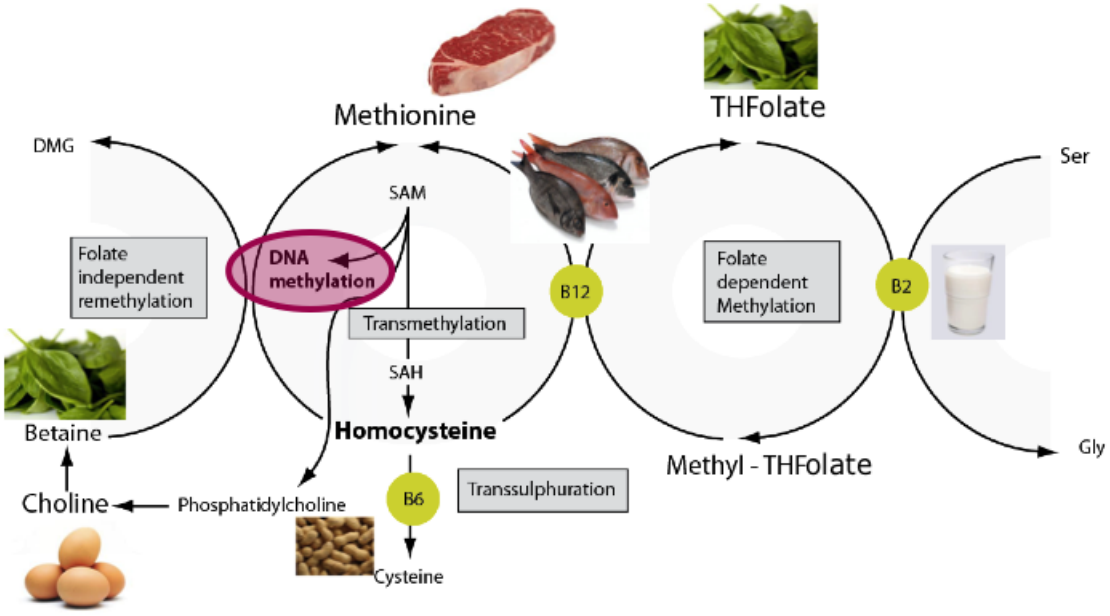
DNA methylation

DNA methylation is the best-characterised epigenetic mechanism to date, due to its chemical stability and the development of sophisticated analytical methods to study it in vitro. In mammals, DNA methylation occurs mainly at C5 of the pyrimidine ring of cytosine bases within CpG dinucleotides, converting it to 5-methylcytosine with the addition of a methyl-group (-CH3). This covalent modification inhibits the affinity of methylation-sensitive DNA binding proteins, affects chromatin structure and ultimately correlates with transcriptional silencing (22).

In normal mammalian cells most of the genomic CpG sites (90-98%) are methylated, and this includes exons, intergenic DNA and transposons (23). Transposons are mobile genetic elements that can 'move' to different locations in the genome. Although the reason is unclear, it is thought that it could be a protective mechanism against expression of unwanted genes (24). Conversely, CpG-islands have a high density of CpG and are found in 50-60% of gene promoter regions (25). CpG-islands are typically unmethylated and act as a regulation switch. A gene can be hypermethylated overall but

have a hypomethylated promoter region which will control its expression (23). One of the 'Human Epigenome project' (HEP) aims is to analyse the patterns of DNA methylation in the regulatory regions of all known genes (19). This is a major undertaking, as an individual's epigenotype is unique. Although some methylation sites are likely to be common for every cell (i.e. those encoding "housekeeping" proteins), distinct tissue-specific methylation patterns exist for many genes with specialised functions in each individual (26).

Figure 2.1: One-carbon metabolism



DMG, dimethylglycine; Gly, glycine; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine; Ser, serine; THF, tetrahydrofolate.

The methyl-groups for DNA methylation are supplied by S-adenosyl-methionine (SAM) (see **Figure 2.1**), and are catalysed by various types of DNA methyltransferases (DNMTs). De novo methylation patterns during gametogenesis, embryogenesis and tissue differentiation are established mainly by DNMT3A and DNMT3B, activity of which is down-regulated upon differentiation of embryonic stem cells and remains low during adult life. Maintenance of these patterns in somatic cells is dependent on different DNMT1

variants, which also have some de novo activity (27). The way these modifications remain stable in cells relies upon the 10-fold higher enzymatic avidity of DNMT1 for 'hemimethylated' cytosine (from DNA where only one of the two strands are methylated). Therefore, when DNMT1 reaches a hemimethylated portion of DNA, with a methyl-group in one of the two strands, it loads the methylation mark onto the newly synthesised strand, as found in the parental strand (27). The accuracy to copy DNA methylation patterns is estimated at ~96% (28). Recent studies have suggested that DNMTs can have a dual role both in methylation and demethylation (29, 30). Studies interfering with the activity of these essential DNMTs during embryonic development have led to abnormal development in humans (31), and genome-wide hypomethylation and embryo death in mice (32, 33).

Histone-based epigenetic marks

Nuclear DNA is organised in the form of chromatin, a highly compact nucleoprotein complex whose basic unit is the nucleosome, which can adopt different three-dimensional conformations (34). The nucleosome consists of a 146 bp stretch of DNA wrapped around a histone octamer (2 pairs of H2A, H2B, H3 and H4 histone proteins). Each histone has a N-terminal tail rich in basic amino acids, which is key to the folding of nucleosomes and inter-nucleosomal interactions (34). Thus, chromatin can appear in two states depending on the necessity for accessibility by transcription binding proteins: euchromatin, more open and transcriptionally competent, and heterochromatin, condensed and transcriptionally silent.

Some genomic regions are made up entirely of heterochromatin, but other regions can undergo transition between the open and the compact conformation (22). Post-transcriptional covalent modifications such as methylation of lysines and arginines, phosphorylation of serines and threonines, ubiquitination and sumoylation of lysines,

acetylation or ADP-ribosylation, mainly at the N-terminal tail of aminoacids and catalysed by specific enzymes (e.g. HAT-histone acetyltransferases-, HMT-histone methyltransfersases-, etc.), can affect transcription by modifying the configuration of the chromatin towards one or the other state. Additionally, histone modifications can affect the activity of non-histone proteins. The effect depends on the type of modification, their number (i.e. diacetylation, trimethylation, etc.) and location (34) within the histone. For example, lysine acetylation is generally associated with increased transcriptional activity (35, 36). Lysine methylation can lead to activation or repression of transcriptional activity, depending on the residue position and the degree of methylation (37). These modifications also play a role in DNA repair and can be reversed by specific enzymes. The existence of these permutations and the way they influence each other and DNA methylation in dictating locus specific transcriptional competence of each cellular type suggests the idea of a "histone code" (22). However, this code has not yet been deciphered fully, as novel modifications are continually being characterised (34). Additionally, as histones completely detach from the DNA during replication, it remains unclear how these patterns are maintained through cellular division. Nevertheless, there is strong evidence of the heritability of histone modifications, and is thought to rely mainly on the (H3-H4)2 tetramer (38).

RNA-based regulation marks; non-coding RNAs and antisense RNAs

Non-coding short RNAs like micro-RNAs (miRNAs) may also trigger post-transcriptional silencing of target genes (39), by specific base pairing with complementary mRNA sequences in the 3' untranslated regions resulting in either inhibition or degradation of mRNA (40). About 60% of human mRNA are targeted by miRNAs and they are involved in cell differentiation, proliferation and apoptosis (41). They can be altered by miss-expression or mutation (22). MiRNAs also have tissue specific expression

contributing to cell identity (42). Antisense RNA was the first non-coding RNA found to have post-transcriptional regulation function.

Non-histone protein-based regulation marks

Although not directly participating in chromatin modifications, three types of proteins are critical for epigenetic regulation: chromatin remodelling complexes, effector binding-proteins and insulator proteins (25). Specific CpG-binding proteins such as MeCP2 interpret DNA methylation by recruiting transcribing or repressing complexes that decipher the histone/DNA methylation marks. Polycomb-group proteins can remodel chromatin and target human gene promoters, throughout development (43). Other factors are involved in chromatin structure and gene regulation, such as nucleosome positioning, whose presence or loss at the transcription start site results in inactivation or activation respectively, maybe by affecting the access for binding of transcription factors (44).

Relations between epigenetic mechanisms

The different mechanisms seem to 'crosstalk' in an orchestrated manner, creating an 'epigenetic landscape' (37) that reinforces the desired signal and establishes a transcriptionally favourable/unfavourable chromatin conformation (19). For example, some specific histone modifications can promote DNA methylation and vice versa (45, 46). Methyl-CpG-binding protein MeCP2, when attached to methylated cytosines, appears to attract other proteins with enzymatic activity that promote nearby histone deacetylation (31). MiRNAs and small interfering RNAs (siRNAs) have also been shown to crosstalk with other epigenetic processes, by regulating the involved enzymes in DNA methylation and histone modification (47, 48). It has been suggested that CpG-islands not associated with annotated promoters could mark the transcription start site of non-coding RNAs (35).

The rules governing the establishment and maintenance of epigenetic marks are still poorly understood, but it is believed that genetic variation in the underlying DNA sequence itself, as well as in the sequence of unlinked modifier loci do play a role (49). A 3-generation study of 48 families observed considerable interindividual variation at methylation of specific genes (i.e. IGF2/H19 and IGF2R) but also familial clustering, suggesting a genetic component to the epigenetic variation (50). However, as discussed in the next sections, environmental factors and stochastical processes are also known to be involved.

2.2.1.2. Epigenetics in development

Epigenetic regulation has been found to operate very actively in normal mammal cell functioning, from conception to aging and death (35). A well-known example of physiological epigenetic activity is genomic imprinting, mono-allelically expressed genes for which only maternal or paternal allele are expressed. Approximately 70 imprinted genes are known in man and over 100 in mice (45). A good example is the insulin-like growth factor-2 gene (IGF2) on chromosome 11p15.5, encoding a developmental growth factor. It is maternally suppressed and paternally activated (maternally imprinted). Many imprinted genes are involved in growth and metabolism. This mechanism is thought to respond to conflicting interests of the parental genomes in the offspring growth and development (paternal imprinting would favour the production of larger offspring, and maternal imprinting favours smaller offspring) (51). Another classic example is X chromosome inactivation in women, where one of the two copies of the X chromosome randomly selected within each cell (52) is silenced by promoter methylation, resulting also in mono-allelic gene expression (53).

Prenatal development

Reprogramming of DNA methylation of the totipotent zygote takes place in primordial germ cells and during pre-implantation (25). Mammal sperm and egg genomes are highly methylated when compared with somatic cells. However, a few hours after fertilisation, rapid demethylation of the paternal DNA takes place by active but yet undefined mechanisms, in addition to histone modification acquisition (54). The maternal DNA follows a slower and more passive process than sperm, with demethylation by simple dilution of DNA methylation during replication, preventing DNMT activity at the replication fork (55). Despite genome-wide hypomethylation, some imprinted regions maintain their methylation status at this point, guided by histone modifications (54). Up to the morula stage, DNA methylation remains reduced and cells are pluripotential with all genes potentially active (45). Simultaneously, primordial germ cells undergo changes in histones and reorganization of chromatin. After implantation genome-wide resetting of epigenetic marks occurs for most of the genome in a lineage-specific manner, and continues to a lesser extent over the rest of the fetal development. Re-methylation at this stage varies upon the part of the embryonic tissue concerned and whereas ectoderm and mesoderm become hypermethylated, primary endoderm and trophoblast remain hypomethylated. There seems to be a sequence of re-methylation dictating the structure and function of each somatic tissue (56).

Inherited paternal/maternal imprinting is erased during gametogenesis to establish new imprints according to the sex of the parent-to-be in the embryonic gametes (57). As the embryo develops into one of the sexes, gene imprinting is placed accordingly (e.g. paternal imprinting during sperm production) (25). DNA methylation is not the only epigenetic mechanism involved in imprinting but seems to be the most important.

Early embryogenesis is critical in the establishment of epigenetic patterns but the timing of events remains unclear in humans. In mice, primordial germ cells appear at day

7.5 of gestation (58), and their epigenetic patterns are completely removed by day 15.5. Methylation patterns are then re-established, with a sex-specific timing. In the male, de novo methylation begins at day 15.5-18.5 of gestation and continues post-natally (59).

Total methylation is thought to change over time (60, 61), whereas methylation at specific loci, such as imprinted genes, seems to be more stable. Thus, most studies so far have concentrated on the more than 70 imprinted genes identified to date in human (57) (web catalogue at http://igc.otago.ac.nz/home.html). Failure to adequately propagate epigenetic marks onto newly synthesised chromatin during cell replication may lead to severe mitotic defects and apoptosis (62, 63).

Cell differentiation

Differentiated cells in a multi-cellular organism are genetically identical, with the exception of B and T lymphocytes. These cells can differ structurally and functionally, expressing only the genes that are necessary for their own activity (54). Epigenetic mechanisms contribute to loss of pluripotency and confer individual cell characteristics, combining genotype, developmental history and environmental influences (54). Tissue-specific CpG-islands methylation of genes essential for development has been found, suggesting programming of DNA methylation patterns (64). Thanks to the activity of enzymes such as DNMT1, tissue patterns are maintained during mitosis (65). Epigenetic marks can be kept through life, until maintenance mechanism start to deteriorate due to oxidative stress or ageing. However, changes can occur also as a response to environmental, behaviour, physiological or pathological signals (18).

Role of environmental signals

A molecular response to environmental changes is an interesting feature of epigenetics. It is considered as a possible interface between genes and the environment in determining phenotype (21). External factors leading to epigenetic dysregulation include dietary excess or deficiency of key nutrients (e.g. folate deficiency), chemical carcinogens such as tobacco smoke or dietary contaminants (e.g. aflatoxin B1), environmental agents (e.g. arsenic), ionising or UV radiation, lifestyle (e.g. stress) or infectious agents (e.g. Helicobacter pylori, human hepatitis virus).(66).

Figure 2.2: Maternal diet-induced hypermethylation at agouti viable yellow (A^{vy}) allele

Effects of maternal dietary methyl supplementation during pregnancy on the epigenome of viable yellow agouti (A^{vy}) mice, leading to variation in coat-color phenotype of A^{vy}/a offspring in isogenic A^{vy}/a animals. The A^{vy} alleles of yellow mice are hypomethylated, allowing maximal ectopic agouti expression. A^{vy} hypermethylation silences ectopic agouti expression recapitulating the agouti phenotype (Waterland *et al*., Mol Cell Biol, 2003)

In mouse models some genes, namely metastable epialleles sensitive to environment influences, have shown to undergo epigenetic changes and exhibit variable expression in an isogenic context. The most classic example of metastable epialleles is the agouti viable yellow (Avy) allele. The Avy allele is a dominant mutation of the murine agouti (A) locus, caused by the insertion of a retrotransposon which acts as an alternative

promoter for this gene which, when active, leads to expression of a yellow coat colour. When the promoter is silenced by methylation, a pseudoagouti coat colour is produced (darker) (see **Figure 2.2**). It has been shown experimentally that maternal supplementation with methyl-donors alters the proportions of the litter coat colour towards pseudoagouti and that this is also associated with higher level of methylation (67-69). Another interesting study showed that the dam's behaviour during lactation (licking and grooming) could also alter DNA methylation of a glucocorticoid receptor gene promoter in the hippocampus in their litter and lead to modified responses to stress (70).

2.2.1.3. The role of epigenetics in disease

The well studied example of the Avy yellow mouse also shows the interesting phenotypical traits of obesity and insulin resistance due to epigenetic silencing, whereas pseudoagouti mice remain lean throughout their lifetime (21). In humans, one of the most useful models for the study of epigenetics and disease is the one provided by pairs of monozygotic twins, of identical genotype, whilst not being phenotypical identical, and exhibiting discordant susceptibility to complex diseases (diabetes mellitus, schizophrenia, cancer, etc.). It has been reported that monozygotic twins have similar DNA methylation patterns at early age, these patterns diverging over the lifetime on overall content and distribution of methylated cytosines and histone acetylation. Many of the discrepancies seemed to refer to Alu sequences (short retrotransposon elements that have the ability to cause epigenetic dysregulation of neighbouring genes, and thus with potential effect in human disease) (60).

Aberrant epigenetic marks have shown to be involved in a whole array of diseases, including diseases evident at birth or in infancy and chronic diseases becoming symptomatic later in life. Although epigenetic marks are deposited early in development,

adaptations occurring through life can lead to major epigenetic reprogramming and thus to disease.

2.2.1.4. Epigenetic technologies

Highly specialised animal models and combination of novel analysis technologies have helped rapidly expanding knowledge in epigenetics, and particularly DNA methylation. Sodium bisulfite treatment of DNA converts unmethylated cytosine residues into uracil, whilst methylated cytosine residues remain unchanged, thus used to identify methylated cytosines and assess DNA methylation level (71). Other techniques exist to identify methylation sites such as methylation-sensitive restriction enzymes, Aza-labelling, immunoprecipitation, etc. Sequencing allows the DNA methylation state to be determined.

Genome-wide profiling with restriction landmark genomic scanning has been successful in identifying differentially methylated sequences, but is labour intensive and associated with bioinformatics problems (72). Oligonucleotide tiling microarrays offer a more affordable alternative though they are being challenged already by next-generation sequencing platforms.

Modern 'Methylation chips' – akin to genotyping chips - allow for a comprehensive genome-wide profiling of DNA-methylation using 'BeadChip' technology, that includes bisulfite treatment, whole-genome DNA amplification, hybridization and single-base extension, fluorescence staining and scanning of chip. The number of CpGs sites that can be interrogated simultaneously, keeps expanding with each new generation of methylation chips and allows for a high-throughput of samples at ever decreasing costs with advancements in technology.

2.2.1.5. Outlook

Many of the difficulties and challenges faced at present by epigenetics have already been highlighted. Ultimately our aim is to fully understand the role of epigenetics in complex diseases for molecular diagnostic and therapeutic applications in medicine. This includes regulation and crosstalk of the epigenetic pathways, the stability and rewriting over time, the extent of external exposure required or the timing of events, particularly whether responsiveness to diet and other factors during development is restricted to more vulnerable and sensitive periods or expanded during the life-course.

To achieve this, it is necessary to develop data resources analogous to those currently existing in genetics, to establish epigenomic maps describing patterns of inter-individual variation. In this sense, the HEP is orders of magnitude larger than the analogous HGP. DNA methylation might be the easiest candidate for epidemiological studies due to the covalent nature of the bond, not being lost during routine DNA extraction, as opposed to the more complex histone modifications. Methylation analysis will also allow for the exploitation of existing DNA biobanks for epigenome mapping purposes (73). This approach must keep in mind that biobanks mostly contain DNA from blood, which will not be a reflection of activity in other tissues and the data should be viewed with caution. Research on epigenetics is challenging but the potential for discovery and for progress in health and disease makes it an exciting field to work in!

2.2.2. RESEARCH PAPER I –' Maternal nutritional status, one-carbon metabolism and offspring DNA methylation: a review of current evidence in human subjects'.

Article cover sheet:

1. Research paper already published

   1.1. Where was the work published? Proceedings of the Nutrition Society

   1.2. When was the work published? February 2012

      1.2.1. If the work was published prior to registration for your research degree, give a brief rationale for its inclusion N/A

      1.2.2.Was the work subject to academic peer review? Yes

      1.2.3.Have you retained the copyright for the work? Yes

         If yes, attach evidence of retention See Appendix II

2. 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 was selected in a postgraduate competition to present preliminary results from my PhD at the Postgraduate Symposium of the Nutrition Society Summer Meeting 2011 and to have a review paper on my research topic published in the Proceedings of the Nutrition Society journal. I wrote the manuscript and my co-authors provided comments which were incorporated prior to submission. The paper was also revised based on the peer reviewer's comments.

Candidate's signature _____

Supervisor or senior author's signature to confirm role as stated in (2) _____

2.2.3. Additional discussion on human studies

The Dutch Hunger Winter cohort produced the first evidence of DNA methylation response to 'in utero' nutritional influences (74, 75). Heijmans, Tobi *et al*. focused initially on *IGF2* DNA methylation, showing an association of prenatal famine exposure with persistent hypomethylation in the offspring, compared to non-exposed. They identified for the first time a critical window in early pregnancy, consistent with the hypothesis of sensitivity during the resetting of the methylation patterns after erasure. However, the changes in *IGF2* methylation could not be related to gene expression and, although they hypothesised the DNA methylation changes might be related to a deficiency in methyl-donors such as methionine, they could not conclude it, nor could rule out other factors such as stress (74). They then expanded this research to a larger number of imprinted and non-imprinted genes within pathways important for growth and metabolic disease. They investigated whether the effect of prenatal famine observed was extended to DNA methylation of other genes and if it mirrored the sex and time specific associations observed through epidemiological studies of prediction of phenotypes in this cohort. The second study concluded that exposure to famine could result in higher or lower DNA methylation in different genes, which suggested a more 'intelligent adaptive response' rather than a 'deficiency response'. An interaction of DNA methylation at certain genes with sex and an effect of maternal malnutrition exposures later in gestation were found to be significant in this study (75). Important limitations of these studies were that: i) they were retrospective and could not link the DNA methylation to the individual intake (even if the ration distribution had been the same, there could be important differences in the actual intakes); and ii) they had not been related to any phenotypic outcome and thus it remained unknown what the real effects of that magnitude of DNA methylation change were. The results of both studies can neither be easily explained by deficiency in methyl-donors nor explain differences in susceptibility to disease at later ages observed in that

population. The authors also tried to link DNA methylation at four of the studied genes with being born small for gestational age, without any association identified (76). However, a later study provided for the first time evidence in humans that increase in infant *IGF2* DNA methylation was associated with maternal use of folic acid and with maternal SAM level as well as with a decrease in birth weight (77). Another study looking at determinants of variation in DNA methylation patterns in newborn infants also highlighted the association of DNA methylation with maternal B12 but not with folate (78). These studies strengthen the case for DNA methylation as a mechanism for fetal programming and for plasticity being dependant on methyl donor supply. However, they do not take account of the complexity of the one-carbon metabolism and the involvement of other components of the diet and substrates of one-carbon metabolism, which might result in non-linear associations. This complexity and the small to modest sample sizes might lead to inconsistencies found between studies in the associations with biomarkers. Also, said studies have the problem of inter-tissue variability of DNA methylation within an individual, which might affect the interpretation. Furthermore, the studies do not (as yet) provide information of phenotypes in the long-term. The Gambian study (79) was new in that it looked at MEs (with no inter-tissue DNA methylation variation by definition), and surprisingly showed a robust consistency in the response to season of conception (hypothesised to be mediated by methyl-donors deficiency) of different genes and a greater effect size than the previous studies in humans had shown. The relevance of methylation patterns at MEs in terms of health outcomes is yet to be elucidated, although animal models show that MEs have the potential to be involved in disease (69).

All in all, the evidence available in humans for maternal methyl-donor deficiency effects on offspring DNA methylation at the design of this study was scarce, but of great interest to warrant further investigation.

2.3.STUDY CONTEXT: WEST KIANG, THE GAMBIA

2.3.1. The Gambia

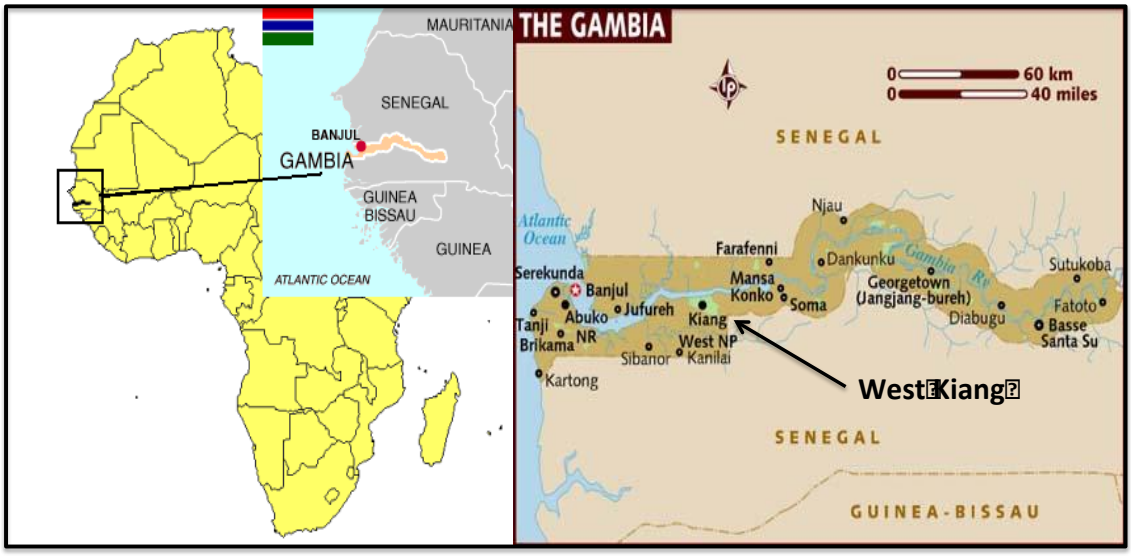
The research presented in this thesis was conducted in the rural West Kiang region of The Gambia. With 11,295 km^2 distributed over a narrow strip of about 15 km on each side of the river Gambia, the Republic of The Gambia is the smallest country on mainland Africa (see **Figure 2.3**, left side). It is located on the West coast, 13º 28' North of the Equator, and bordered to the North, East and South by Senegal. The Gambia gained its independence from the UK in 1965.

The Gambian population, of five main ethnic groups (42% Mandinka, 18% Fula, 16% Wolof, 10% Jola and 9% Serahuli) (80) and predominantly Muslims, was estimated at 1,73 million in 2010 (81). The Gambian 2010 Human Development Index was 0.390, placing it in the 151th position out 169, with a Gross National Income per capita of $1,358, a 34.3% of the population in severe poverty (below poverty line) and a life expectancy at birth of 56.6 years (82). According to the 2011 UNICEF report on The State of the World's Children, the prevalence of HIV in 2009 was 2.0%, the under-five mortality rate was 103 per 1000 and the proportion of infants with low birth weight was 20% (83). The rural population included 63% of the total inhabitants and the adult literacy rate was 45% (83).

The Gambian population, as happens in other sub-Saharan countries, is in the early stages of shifting towards urbanisation and undergoing a nutrition transition (84). Both these demographic and nutrition transitions are likely to lead to changes in lifestyle and a 'double-burden of disease', where malnutrition and infectious diseases coexist with an increasing prevalence of non-communicable diseases (NCDs), including obesity, cardiovascular disease or type -2 diabetes (5). Hypertension is prevalent in urban Gambia and not rare in some rural areas of the country (85). Obesity among Gambian women 35

years or older was found to be between 32.6% in urban areas, down to the 4% in the rural area around Farafenni (86).

Figure 2.3: Map of The Gambia, West Africa

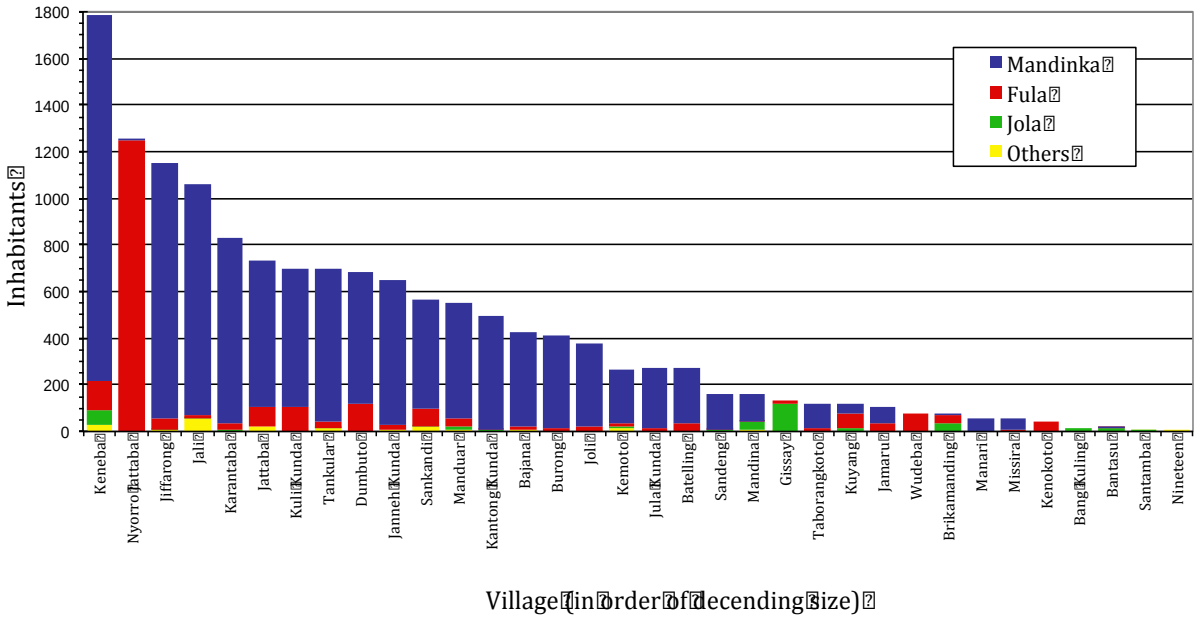


Reproduced from The Gambian Experience (regional map) and The Lonely Planet (Gambian map)

2.3.2. West Kiang

West Kiang is a district in the Lower River Region (see **Figure 2.3**, right side) with 13,800 inhabitants registered in 2009 (West Kiang Demographic Surveillance System, WK DSS) in 34 villages, in approximately 750 km^2 with limited road access. Of these, 2,480 were estimated to be women of reproductive age. The population is predominantly Mandinka (see **Figure 2.4**). Villages in West Kiang are divided into compounds, where extended families live together, with an average size of 16 people per compound (WK DSS). Roles for men and women are distinctive and well defined. Women are the main workforce: cooking, growing and harvesting the staple foods (87). Men are also involved in agriculture, but mainly in groundnut cultivation (traditionally the main cash crop). Rural subsistence farming is the main livelihood. Thus, income and eating patterns fluctuate strongly according to the annual farming calendar, heavily influenced by the monomodal annual rainy season (88).

Figure 2.4: Village population size by ethnicity in West Kiang



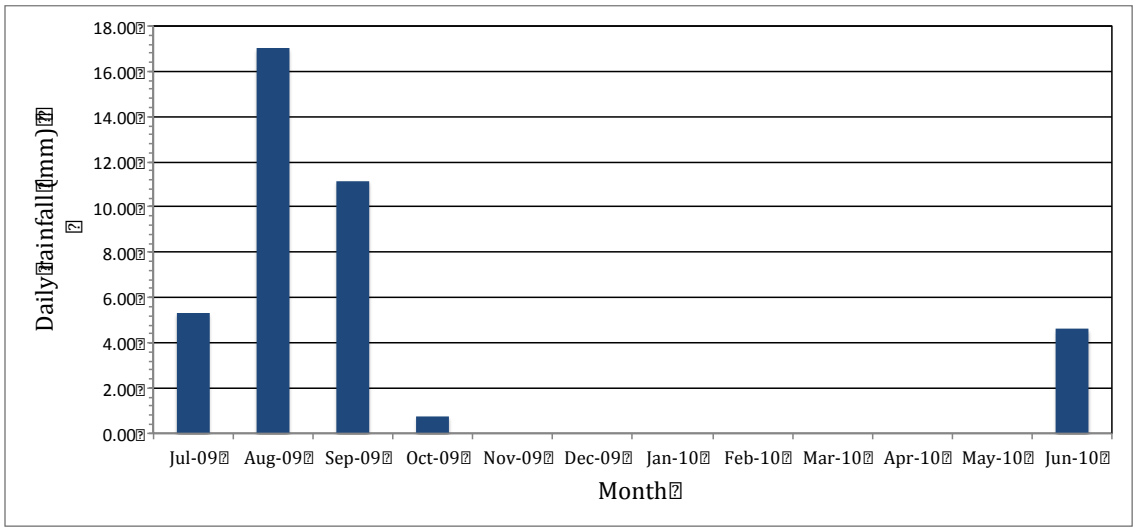
Source: West Kiang Demographic Surveillance System, data from 2010 (provided by T.Fulford and B.Hennig)

2.3.3. Seasonality and seasonal findings

Agricultural calendar

The tropical climate in The Gambia has two very distinct seasons, a hot rainy season from June through November, and a cooler dry season, covering the remaining months. The total annual rainfall monthly recorded in Keneba between 2009 and 2010 is detailed in **Figure 2.5**. This seasonality strongly shapes the daily life in West-Kiang. The cereal crops are harvested between September and December. Thus, food supplies are particularly plentiful in November and December, while in the rainy season, particularly in July and August, food shortages often occur, as food stores from previous year's harvest and cash are lower (88). At this time, there is also a high agricultural workload, undertaken predominantly by women. This can lead to a negative energy balance (89), and for this reason this season has traditionally be known as 'hungry' season. November to May is conversely termed 'harvest' season.

Figure 2.5: Daily rainfall (mm) monthly average in Keneba from June 2009 to September 2010



Data obtained from the Water Resources Department of the Gambian Government

Dietary patterns

The effect of seasonality in the eating patterns and intake is both quantitative and qualitative. In addition, the agricultural work also limits available time for cooking, which further affects the types of food that can be prepared during the farming (rainy) season (90).

Between January and June all the cereals and groundnuts (*Arachis hypogaea*) are available, whilst the rainy season is characterised by low stocks of both. The main cereal crops produced are rice (*Oryza sativa*), millet (*Pennisetum typhoideum* and *Pennisetum gambiense*), maize (*Zea mays*) and sorghum (*Sorghum margaritiferum*). Rice is the staple of choice and is eaten with main meals or in porridge. Porridge is very popular as breakfast and there is a different range of types, from a thinner type ('mono') to the thickest one (`sato´). Boiled rice is eaten with different types of sauces. The other staples are often consumed steamed ('futo') (88). Sauces normally include a combination of oil, groundnuts, green leaves ('jambo'), vegetables and dry or fresh fish. Cooking is always done by women, either in small kitchens build of mud bricks with little to no ventilation or

outside the house, using wood as fuel. Meals are usually shared by several family members and are eaten from the same bowl.

Groundnuts are used as 'cash crops' and also eaten as a good source of protein (including methionine). Groundnuts sauces are highly appreciated but, during the rainy season shortage other less preferred foods can be consumed as a replacement. In this period, bush foods such as green leaves and fruits, are plentiful. This can potentially lead to more dietary diversity. During the dry season the availability of wild foods and leaf vegetable decreases, but usually home-grown vegetables (e.g. tomato, okra, etc.) are available from home gardens by April/May (91). Some leaves such as cassava leaves are still available, but leaf sauces are not frequently eaten during this period, because other preferred foods are plentiful. Shortages begin to be felt in June (91).

During the dry season, there is higher availability of cash for buying items such as oil, meat or fish (88). Meat consumption is generally low, and mainly restricted to special occasions and celebrations. Fish, however, is easily available and frequent in the rural West Kiang diet all over the year. In the river Gambia, several types of fish may be caught (i.e. bonga –oily fish, catfish, barracuda, tongue-sole and shrimps).

It is believed that some changes in dietary patterns have occurred over the last decades, particularly increased palm/vegetal oil usage in cooking (90). Traditionally, red palm oil was only purchased by the more wealthy families (92). According to a focus group study conducted recently in West Kiang, the food choices are dependent not only on individual preferences but also on cash availability and access (90). Nowadays, the food availability might have been modified, at least in part, by the increasing cash in the region from national or international 'remittance economy', where family members who have migrated to urban areas or abroad send money back to rural areas (93). Improved transport in the region might also play a role. At present, increased influence from the coast and 'imported' products (e.g. tinned milk, sardines, candies, biscuits or mayonnaise)

exist in the area and this might have contributed to changes in the dietary practices. Yet, the population as a whole still has a limited consumption of this type of foods and remains largely dependent on the locally grown crops (personal observation).

Nutrition

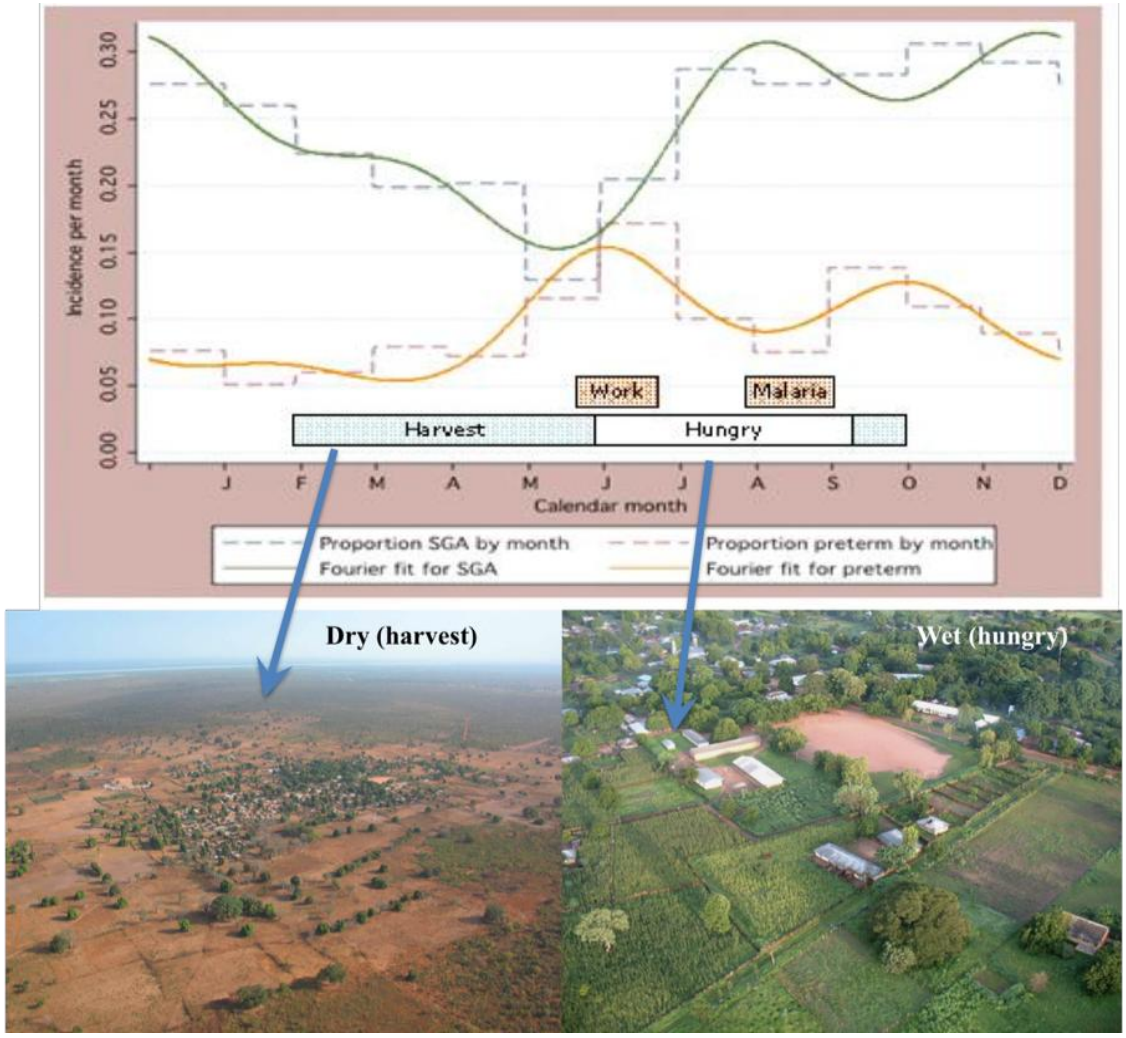
Seasonal variations in availability of leaves, cereals and groundnuts, result in nutrient intake variation. The population of several villages in West Kiang (core villages) has been well characterised in this respect: dietary studies have been conducted for the last 40 years, particularly in pregnant and lactating women. Such studies specifically focused on energy intake (94) and also in macro- and micronutrients intake and nutritional status, including calcium, phosphorus, folate, vitamin C and vitamin A (91, 95). These studies have clearly shown seasonal variation in different nutrients, with specific patterns for each nutrient. For instance, the rainy season was associated with increased calcium intakes (by 16%) whilst decreased phosphorus consumption (by 15%) (91). Few of the one-carbon metabolism methyl-donors and cofactors have been investigated. Most information available (B2 and folate) dates back to 1978-80 (95). A specific Gambian food database was produced for the dietary studies (96) but this did not contain information related to methyl-donors and cofactors composition. In 1997, plasma homocysteine, folate and B12 concentrations were studied in women during the dry season. The mean plasma homocysteine concentration (geometric mean 8.92 μmol/l) in Gambian non-pregnant women was similar to that of UK women (97).

Health and reproductive outcomes

Seasonality has also a number of effects in population health that have been thoroughly studied. During the rains an increase on disease transmission has been observed, including falciparum malaria, pneumonia and gastroenteritis, and therefore a peak of prevalence of childhood-prone diseases and malnutrition in September and October (98). Moreover, the risk of infection-related mortality is 10-fold higher in adults born during the nutritionally poorer rainy season, suggesting that immune function could be compromised during early-life development (99). Several signs of immunological deficit such as reduced thymic size and T-cell output have been identified. Thus, the implications of DOHaD are thought to extend beyond NCD of affluence, to those involving susceptibility to infectious diseases (100).

Effects of Gambian seasonality at conception and birth have also been observed on reproductive outcomes, namely prematurity and small for gestational age births (SGA) (101). As observed in **Figure 2.6** prematurity and SGA showed divergent seasonal patterns. SGA was highest at the end of the rainy season, peaking in November at 30.6%, with the nadir of 12.9% in June, inversely to maternal weight changes (101). Peaks in prematurity, in July (17.2%) and October (13.9%) closely paralleled increases in agricultural labour and malaria infections respectively (101). Nutritional supplementation with high-energy groundnut biscuits (4.3 MJ/day) during pregnancy has confirmed the key role of nutrition in this seasonal differences. Prenatal dietary supplementation reduced retardation in intrauterine growth, with a highly significant effect on the average birth weight (102, 103). This effect was considerably greater during the rainy season (increase of 201 g *versus* 94 g), where mothers are most 'at-risk' (103). Also, a substantial reduction in the prevalence of stillbirths and perinatal mortality was observed.

Figure 2.6: Month-by month percentage of SGA and premature infants



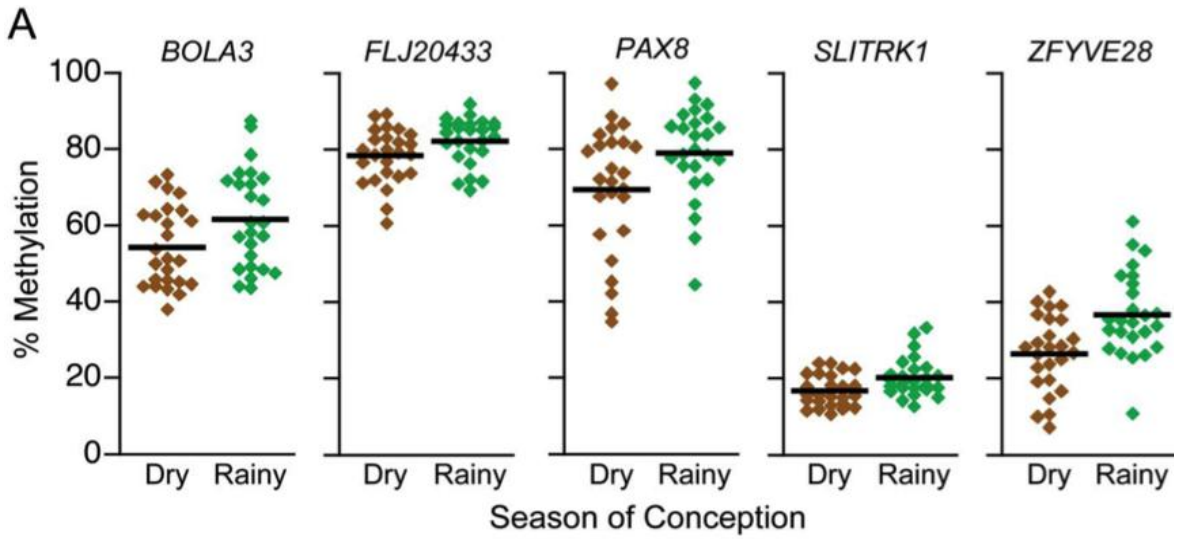
Fitted seasonality models for preterm and small-for-gestational-age (SGA) infants. The dashed green and yellow lines represent the proportion of SGA and preterm infants per calendar month respectively. The solid green and yellow lines represent the Fourier fit for SGA and preterm infants. (Modified from Rayco-Solon *et al.*, Am J Clin Nutr 2005 [101])

2.3.4. Epigenetics

Over the last years, preliminary epigenetic studies have also been conducted in West Kiang. Based on the hypothesis that periconceptional maternal micronutrient supplementation affects fetal DNA methylation, two studies looked at how the epigenome was affected in offspring of women receiving a United Nations International Multiple Micronutrient Preparation (UNIMMAP) supplement during pregnancy. DNA methylation genome-wide within gene promoters (104) and at imprinted genes (105) was explored, in cord blood and postnatal DNA samples of their offspring. The data, although based on small numbers, suggest that periconceptional maternal nutrition in humans can alter offspring methylation at imprinted loci and is an important determinant of newborn whole genome methylation patterns (104) (105). Sex-specific effects of micronutrient supplementation on imprinted genes were also observed (105).

Epigenetics have also been investigated with respect to seasonality in West Kiang, as discussed in Section 2.2.2. DNA methylation in two groups of children conceived either during the dry or the rainy season was compared. DNA methylation at metastable epialleles (i.e. alleles that are variably expressed due to epigenetic modifications, MEs) was higher in children conceived during the rainy season (August-September) than during the dry season (March-May), as shown in **Figure 2.7** (79). This suggested the possibility that methyl-donor intake might be higher during the rainy season and raised the need for further investigation.

Figure 2.7: DNA methylation at putative MEs is influenced by season of conception in The Gambia



MEs: metastable epialleles

Percent methylation at five MEs in DNA of Gambian children, by season of conception. Each diamond represents one individual, and the black lines group means (n = 25/group). At all genomic regions, DNA methylation is higher in individuals conceived in the rainy season (*BOLA3* P = 0.03, *EXD3* P = 0.03, *PAX8* P = 0.02, *SLITRK1* P = 0.006, *ZFYVE28* P = 0.002; overall P = 0.0001) (Waterland *et al*., PLoS 2010 [79]).

2.4.STUDY RATIONALE

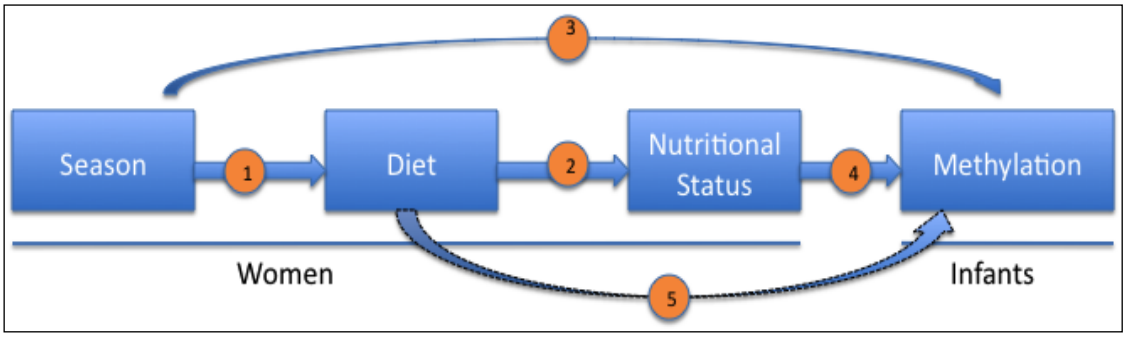
As presented above, the marked pattern of seasonality in this rural Gambian setting provides an excellent research opportunity to investigate effects of maternal nutrition on offspring outcomes. This experiment of 'natural' variation in dietary and other lifestyle pattern enables the testing of hypotheses related to nutritional deprivation, that might be unethical to answer through a traditional intervention in humans. Such questions include a deeper understanding on the mechanisms mediating fetal programming. As previously outlined in the literature review, and based on previous findings from epigenetic studies in West Kiang, DNA methylation is a candidate mechanism worthwhile to pursue due to its key role in in early fetal development, its sensitivity to environmental factors and dependence on dietary factors. The hypothesis of the role of maternal nutrition in epigenetics, tested in animals, has obtained promising results. This thesis expands these observations to study early development factors affecting DNA methylation in humans.

2.5. HYPOTHESIS AND OBJECTIVES

The current study set out to explore the hypotheses that, in rural Gambia, seasonality in food availability leads to differences in biomarkers status for methyl-groups supply and such changes, when occurring around conception can lead to detectable differences in DNA methylation patterns of women's offspring.

The specific objectives are summarised in **Figure 2.8**:

Figure 2.8: Conceptual framework of the study



Note: the encircled numbers correspond to the 5 objectives below

1) To establish the annual patterns in the intake of methyl-donors and cofactors in women of reproductive-age, by assessing the daily intake and analysing the nutritional content in local foods.
2) To assess the annual patterns in the nutritional status for a comprehensive set of methyl-related intermediary metabolites.
3) To replicate the effect of seasonality in the DNA methylation of infants born to women who conceived at the peaks of the rainy/dry seasons.
4) To assess whether seasonal differences in offspring DNA methylation can be mediated by maternal methyl-donor status, by exploring their relationship .

5) Finally, to assess the relationship between diet and nutritional status and thus (indirectly) the relationship between maternal diet and DNA methylation outcomes in their infants.

To achieve these aims and objectives, the research projects described hereafter were designed. These objectives were addressed through two separate but integrated studies running in parallel, as described in detail in the next Chapters. The nature of seasonality in The Gambia lent itself for an 'experiment of nature', as defined by the recent MRC guidance for the use of natural experiments to evaluate population health interventions 2012 (106, 107). These guidelines were produced after the present study was designed and conducted, but the potential limitations highlighted by the document are acknowledged. Even if this approach needs special care in the interpretation of findings (i.e. bias and confounding issues) and limits the possibility of establishing causal inferences, the design proposed will allow for exploratory work of exposures that would be unethical to manipulate without further knowledge and may provide hints on potential causation (106).

CHAPTER 3: OVERALL STUDY DESIGN AND METHODS

3.1.FIELD SITE

The MRC has been involved with work in West Kiang, district of The Gambia, since the late 1940s. Since 1974, the MRC has run a permanent field-station in Keneba (MRC Keneba, Appendix III (A)), which is the largest village, in the centre of West Kiang. MRC Keneba has a well-established infrastructure and receives logistical support from MRC Fajara, the main operational MRC Unit, The Gambia, approximately 160 km away from Keneba on the Atlantic coast. In addition to its international research agenda, in conjunction with the regional health team, the MRC provides medical care to the whole region, focusing primarily on maternal and child health. Furthermore, the MRC is the biggest employer in the region, and the second largest employer in the country. All this has helped develop strong links with the local community, who are very supportive of the MRC's research programmes. The WK DSS was set up in 2004; it covers the whole region and facilitates the characterisation and the tracking of its population. Compound heads are visited four times a year to record births, deaths, in- and out-migration, and other vital events such as marriages. The facilities at MRC Keneba include clinical services, a laboratory for clinical and research analyses, several research rooms with specific equipment, a data office, and a fleet of vehicles to conduct the fieldwork. This provides a good platform for research studies, where accurate information can be collected, samples can be immediately processed and stored, and participants can be followed as part of long-term studies.

3.2.STUDY POPULATION: MAIN GROUP AND INDICATOR GROUP

The study presented in this thesis is an observational prospective proof-of-principle research, conducted across West-Kiang. The primary exposure was season. Maternal dietary intakes were used both as outcome of season and secondary exposure. Biomarkers of methyl-donor status were looked at as proximal endpoints, cross-sectionally to season and dietary intake (see Figure 2.8). Maternal biomarkers were also examined as exposure for the offspring DNA methylation, which was looked at as a distal outcome prospectively of all other variables. Under the term biomarkers, this thesis refers to all the laboratory analysis done in maternal blood (plasma and red blood cells (RBC)) to determine their nutritional status in respect to one-carbon metabolism.

The study was composed of two complementary studies, termed from this point onwards as the 'indicator group' study and 'main group' study. The fieldwork for both studies started in July 2009 and concluded in June 2010 (indicator group) and June 2011 (main group).

The joint Gambian Government/MRC Ethics Committee (SCC/EC 1151) and the LSHTM Ethics Committee (EC 5525) granted ethical approval for all aspects of this work. The standardised information sheets and consent forms are given in Appendix IV and V.

3.2.1. Study 1: Indicator Group

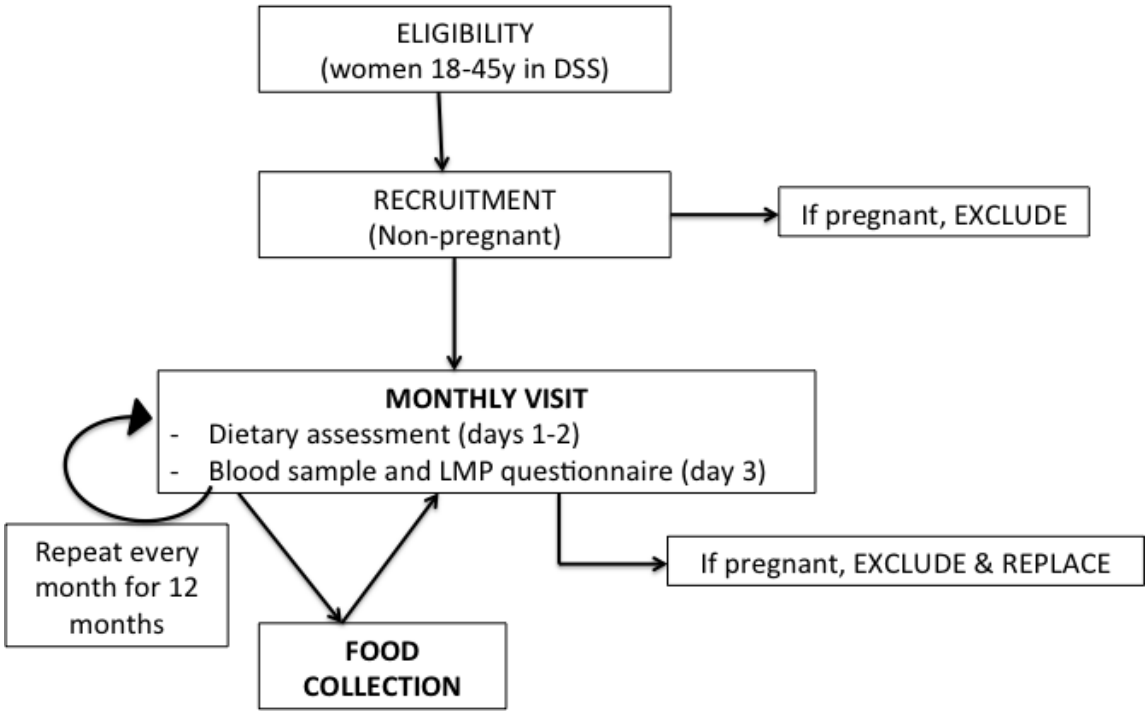
Thirty healthy non-pregnant women of reproductive age were followed up monthly during one year. This indicator group was set up to meet objectives 1 and 2, i.e. provide data on annual variation in the intake of a comprehensive set of methyl-donors and cofactors and a corresponding set of blood biomarkers of one-carbon metabolism. This group also provided information of dietary intake to help estimate seasonal differences in intakes of the women in the main study within the first weeks of pregnancy

and achieve indirectly objective 5 (establish relationship between early pregnancy maternal diet and infant DNA methylation outcomes).

In the indicator group, a detailed 2-day dietary assessment was performed monthly (Appendix III (B)). Following the dietary assessment, blood was collected to assess selected biomarkers and determine associations with diet cross-sectionally (see **Figure 3.1**). The dietary assessment consisted of direct observation of the meal preparation and weighing all ingredients for the cooking to determine the recipe nutritional composition. The participant's intake was then weighed throughout the day. All data were recorded by field assistants, resident in each of the three villages, and recorded. The information from two forms (Recipe form, Appendix VI, and Dietary intake form, Appendix VII) was used to establish a list of Gambian common foods and ingredients, for which samples were collected for composition analysis of vitamins B2, B6, B12, folate, choline, betaine and methionine. Further details are given in Chapter 4, sections 4.1 and 4.2.

Only non-pregnant women of reproductive age were recruited. This was done for two main reasons: (i) our research hypothesis was that extensive demethylation and remethylation of fetal DNA takes place around conception and early development (i.e. when it may not be known yet that they are pregnant). Therefore this group was deemed a good proxy of periconceptional status; and (ii) plasma choline levels, and other components within one-carbon metabolism, may be modified during pregnancy or menopause, as they are up-regulated via oestrogens. Women who became pregnant or self-withdrew from the study were replaced by another eligible woman from the same village (see Section 4.1).

Figure 3.1: Flow diagram of the indicator group set-up and recruitment



DSS: Demographic surveillance system; LMP: Last menstrual period

3.2.2. Study 2: Main group

A cohort of approximately 1600 women of reproductive age, from all 36 villages in West Kiang, was followed-up monthly for pregnancy detection. The objective was to identify women conceiving at the peak of the rainy (July-September) and the dry (February-April) season and to follow-up their infants (see **Figure 3.2**). The decision on the peaks of the two seasons was mainly based on the local rain patterns. As shown in Figure 2.5, July, August and September are the wettest months of the year. The driest months include February to May, but some years rains can start as early as May. The final number of women-infant pairs fully enrolled into the study was limited by the number of full term pregnancies conceived during the two season time windows. The ability for early detection of pregnancy was limited in part by willingness of women to disclose their pregnancy status, thus the cut-off for 'early' blood sample collection was set within the first 16 weeks.

The main group contributed to objectives 2, 3, 4 and 5 (see **Table 3.1**), by determining maternal nutritional blood biomarkers status close to conception and following up their offspring to establish association with DNA methylation patterns.

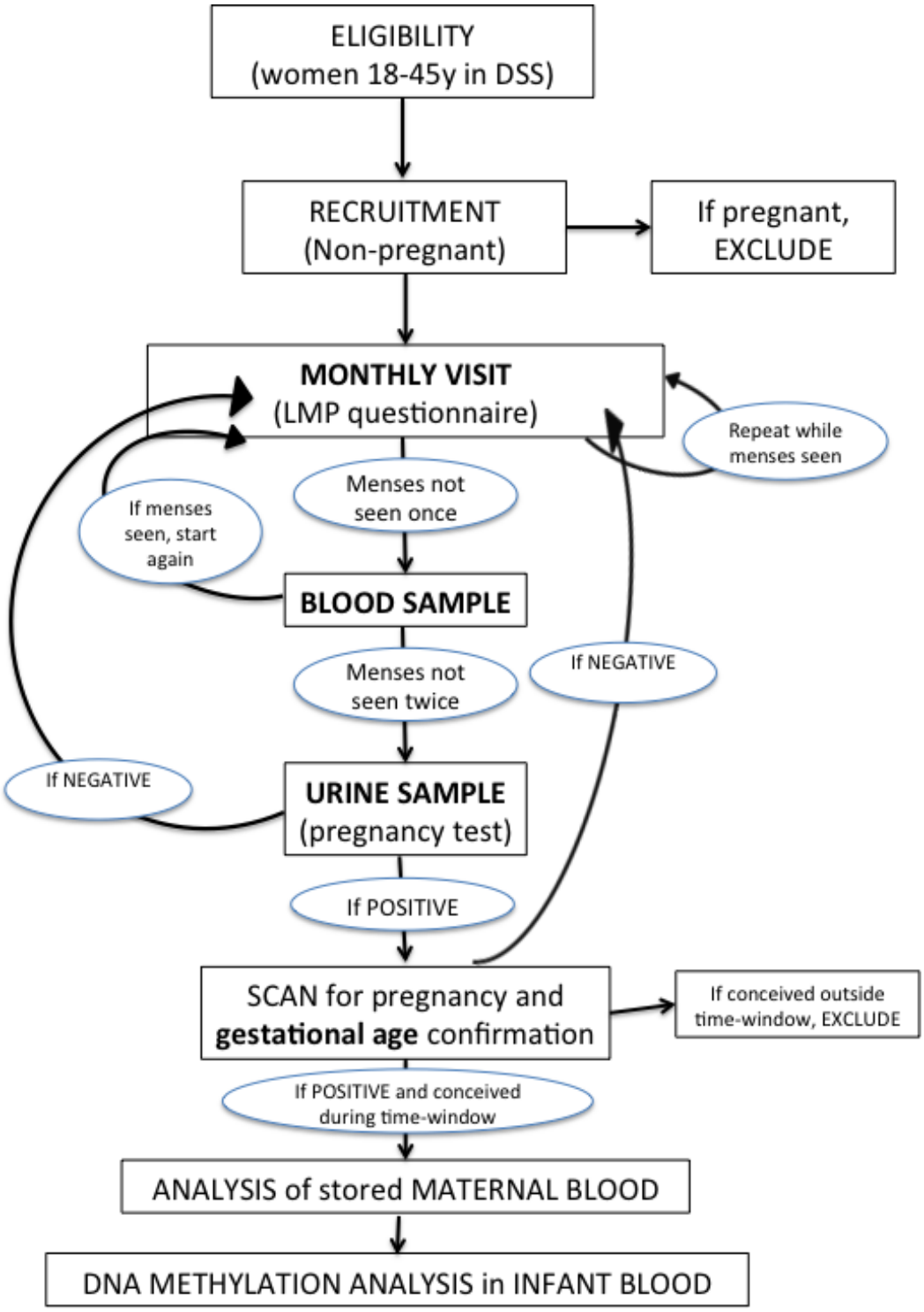
In order to protect the right of women to keep their pregnancy confidential, this study was designed in such way that blood samples were collected prior to pregnancy detection, based on the lack of a menstrual period (as determined by a last menstrual period (LMP) questionnaire, Appendix VIII). To avoid unnecessary blood collection of women unlikely to be pregnant, a decision tree was included in the training of field workers (see **Figure 3.3** to assess the need for or the exemption from blood collection). Further details about the fieldwork are given in Chapter 5, sections 5.1 and 5.2.

The women of the indicator group were excluded from the main group recruitment but were given the opportunity of joining it should they become pregnant. Recruitment into the main group study was done jointly with a trial of Early Nutrition and Immune Development (the ENID Trial) (ISRCTN49285450) examining the effect of pre-natal and infancy daily nutritional supplementation (micronutrients and/or energy) on infant immune development. ENID is a large and long-term project, within the same target population (all women of reproductive age in West Kiang), and this study (Methyl-Donors and EpiGenetics, MDEG) was partially nested within this trial.

This helped maximise the local resources, whilst minimising the burden on the participants, and provided this study with the opportunity of accurate gestational age assessment by obstetric ultrasound. At approximately 12 weeks of pregnancy ±2 weeks, pregnant women were randomly assigned to one of four pre-natal nutrition (micronutrient and/or energy or placebo) supplements. All participants, included those who were not in ENID, were given 60 mg iron and 400 μg folate (FeFol), either as a part of the ENID nutritional supplement, or alone. This is in line with the current national

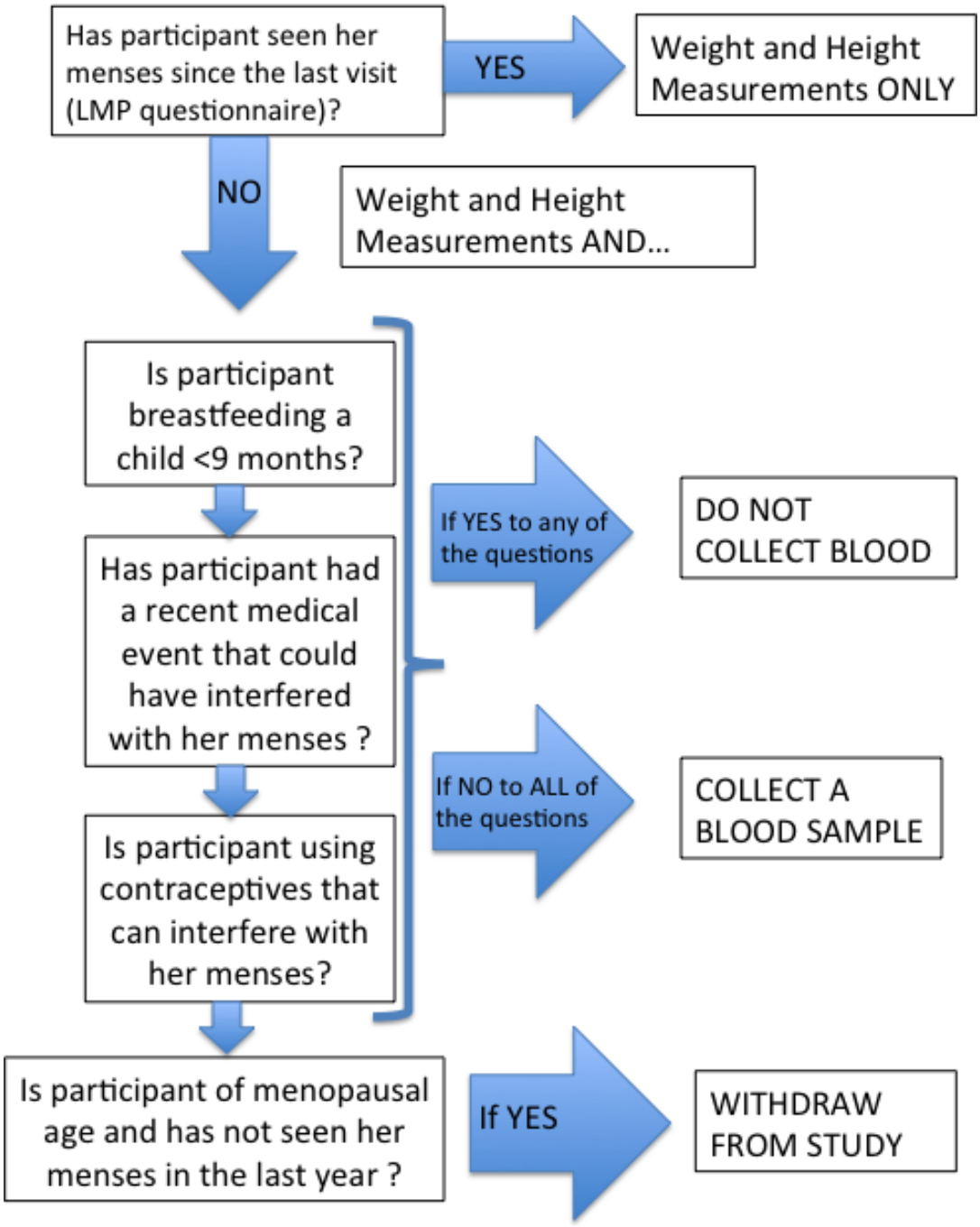
guidelines in The Gambia, which recommend this daily supplement to all pregnant women until the end of pregnancy.

Figure 3.2: Flow diagram of the main study set-up and recruitment



DSS: Demographic surveillance system; LMP: Last menstrual period

Figure 3.3: Decision tree for field blood collection from participants following the report of a missed menses



Assessment done as a part of the monthly LMP questionnaire; LMP: Last menstrual period

3.3.DATA ANALYSIS

3.3.1.Data entry

Specific identification numbers (study IDs) were given to participants of both studies at the time of recruitment, including a check letter. Data entry of the monthly visits for both the indicator and the main group was done real-time in a specifically designed Access database. More details on the entry of dietary intake data is given in Sections 4.1. and 4.2.

3.3.2.Data manipulation

Due to the quality control measurements put in place for the data entry (i.e. automatic range of maximum/minimum 'acceptable' values to prevent entry of erroneous values and double data entry), the need for data cleaning at the end of the study was limited. Once laboratory analyses were complete, checks were done to identify values out of range (where normal data ranges were available) or erroneous (e.g. repeated samples). Data were then plotted (box-plots) in order to check for outliers and any other noticeable errors. Data errors that could not be rectified and clear outliers were dropped from the dataset. Dietary intake data were single-entered and therefore, intensive data cleaning was performed to identify mistakes, as detailed in Section 4.2.2.

Continuous outcome variables and covariates were assessed for normality, graphically by drawing histograms and numerically by applying the Shapiro-Wilk test. Almost all continuous covariates of dietary intake, blood markers and DNA methylation were skewed. Therefore, they were transformed to logarithms (intake and biomarkers) or logits (DNA methylation percentages) to normalise the distributions. Summary statistics of dietary intake and biomarkers were reported as geometric means.

3.3.3. Statistical analysis

All statistical analyses were conducted using Stata 11 (Stata Corporation, College Station, Texas, USA). The analysis was different by study group/objective, as detailed in the following sections. The common link in both study groups was the women's biomarker concentrations in blood. The specific statistical analyses for each study group are described in the respective relevant Chapters (Chapters 4 and 5).

3.3.3.1. Summary of data analysis by objective

The specific details of the statistic analysis for each of the studies are given in Chapters 4 and 5. In summary:

<u>Objectives 1 and 2: assessing the annual patterns in the dietary intakes and blood biomarkers of one-carbon metabolism related substances in West Kiang</u>

The indicator group study was set up to describe and establish annual seasonal trends in methyl-donor and cofactors dietary intake and one-carbon metabolism blood biomarkers. Thus, the analysis of this study group was mainly descriptive. The overall level of dietary intakes and blood biomarkers were assessed and compared with international recommendations/cut-offs of normal range. The data represented repeat observations for each participant on consecutive months, and were thus not independent. This correlation was accounted for using random effect models with Generalised Least Squares (GLS) regressions. Seasonality patterns were modelled using the truncated Fourier series approach. This mathematic model is good to define seasonality, as the Fourier series are smooth linear functions whose terms are approximately orthogonal to one another and inherently cyclic (108). They offer flexibility in the dimensionality of the fitted seasonal effect: truncating the higher-order terms of the series removes higher-frequency noise (108). Generally, fitting the first two pairs of Fourier terms (F2 model: sin,

cos, sin2, cos2) were significant in these data and deemed sufficient. Further details of this analysis are described in Chapter 4. In the main group, biomarkers were compared (one-way analysis of variance, ANOVA) in dry versus rainy season measurements. Biomarkers status differences between seasons were also assessed in the main group with one-way ANOVA. More details are given in section 5.1.

Objective 3: assessing seasonal DNA methylation differences

One-way ANOVA was used also to compare mean DNA methylation percentage at each CpG site, by testing differences in DNA methylation of infants conceived during the rainy and the dry season. Multivariate analysis of variance (MANOVA) was used to assess methylation at each gene, considering all its CpG sites simultaneously. Effect of non-nutritional determinants on DNA methylation was also assessed (e.g. infant sex) by multivariate regression analysis. Further details are given in Section 5.1.

Objective 4: assessing the association of maternal biomarkers with infant DNA methylation

The initial analysis was centred on plasma SAM and the SAM:SAH ratio as 'bottle neck' factors. In order to establish the blood biomarker concentration at conception, back-extrapolation employing data from the indicator group (objective 2) was used, after comparison of seasonal patterns between both groups for reliability and stability. Multiple regression was used to explore the prediction capacity of the different biomarkers of DNA methylation, examining the independent influence of each biomarker on DNA methylation. Their interactions were also examined. Principal component analysis was used to investigate all biomarkers jointly, and multilevel analysis to pool the different loci together. All analysis were adjusted by infant sex, as sex differences in DNA methylation have been repeatedly reported (75, 104, 105). Similarly, maternal age and BMI were

included in the models, as it can affect both infant DNA methylation and maternal biomarker blood concentrations. Further details are given in Section 5.1.

Objective 5: assessing the association of maternal dietary intake with infant DNA methylation

After having established the annual trend for dietary intake and biomarkers in objectives 1 and 2, and having used these to estimate the biomarker blood concentration at conception by back-extrapolation in objective 4, the same approach was planned for the dietary intake information to estimate the maternal intake around conception. The prediction capacity of biomarkers from the dietary intake was assessed to establish the viability of such approach. Details are given in Chapter 4.

Table 3.1: Overview of objectives and methods

Objective	Study group (Paper/Chapter)	Methods
1. To assess the annual patterns in the intake of methyl-donors and cofactors	Indicator group (Paper II/Chapter 4)	Monthly 48-weighed-records throughout a year (2 days x 12 months x 30 women) Food samples for analysis of vitamins B2, B6, B12, folate, choline, betaine, methionine content)
2. To assess the annual patterns in blood biomarkers of a comprehensive set of one-carbon metabolism related metabolites in women of reproductive age, monthly or as close as possible to time of conception	Indicator (Paper II/Chapter 4) & Main group (paper III/Chapter 5)	Monthly blood sample throughout a year (1 sample x 12 months x 30 women) LMP questionnaires monthly One-off blood collection in early pregnancy -at first lost menses (decision tree) Pregnancy test at second lost menses Determination of gestational age by ultrasound
3. To replicate the effect of seasonality at time of conception on the DNA methylation of infants	Main group (Paper III/Chapter 5)	Determination of gestational age with ultrasound scan (confirmation that date of conception fits within pre-defined season window) Back-extrapolation of nutritional status of mother to time of conception using indicator group data DNA methylation analysis of candidate loci by pyrosequencing in offspring by season of conception
4. To test for relationship of these seasonal differences in offspring DNA methylation with the methyl-donor status of their mothers	Main group (Paper III/Chapter 5)	Combining data from objectives 2 and 3
5. To assess (indirectly) the relationship between maternal diet during early development and DNA methylation outcomes in their infants	Main + Indicator group (Chapters 4 and 6)	Combining data from objectives 1-4

3.3.3.2. Sample size calculations and power

Indicator group

Precise sample size/power calculations were not possible due to the scarce information existing on variation in the dietary intake and blood levels of the substances under study. Instead, this study must be regarded as an exploratory study to provide such data. It was judged that sampling 30 women to allow for losses to follow-up, withdrawals and pregnancies, together with a replacement system, at 12 time points evenly spaced across the year, should provide a good estimate of the seasonality with resolution at the level of one month.

Main study

The sample size for the main study group was subject to the number of pregnancies occurring during the study period and willingness of women to participate. Based on numbers from previous studies undertaken in this population, it was considered feasible to recruit a maximum of 100 mother-infants pairs per 3-month season. For the testing of the hypothesis of differences in DNA methylation between infants born in the rainy and dry season, it was established that 100 infants per season would have enabled the detection of 0.46 standard deviations in methylation level with power 90% and at 5% significance.

Part II:
RESEARCH PAPERS WITH RESULTS

CHAPTER 4: INDICATOR GROUP

This Chapter describes in detail and discusses the methods and results of the statistical analysis performed in the indicator group study, investigating the annual variation in dietary intake of methyl-donors and cofactors and blood concentrations of one-carbon metabolites in women of reproductive age. The main methods and findings have already been submitted as a peer-reviewed paper, which is presented as Section 4.1. In addition, this Chapter provides further unpublished information on methods and results and explores and discusses the data in greater detail.

4.1. RESEARCH PAPER II: 'DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and co-factors in rural African women'

Article cover sheet:

1. For a 'research paper' prepared for publication but not yet published

   1.1. Where is the work intended to be published?

   This paper presents the main findings of indicator group study and was submitted to the American Journal of Clinical Nutrition, the accepted modified final version is included as Appendix IX.

   1.2. List the paper's authors in the intended authorship order

   Dominguez-Salas P, Moore SE, Cole D, da Costa KA, Cox SE, Dyer RA, Fulford AJC, Innis SM, Waterland RA, Zeisel SH, Prentice AM and Hennig BJ.

   1.3. Stage of publication – Not yet submitted/Submitted/**Undergoing revision from peer reviewers' comments**/In press

   Submission on 30/07/2012

2. For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

   I conceived and designed the study, conducted the research and analysed the data, in collaboration with some of the authors (Moore, Cox, Fulford, Prentice, Hennig). I wrote the manuscript and my co-authors provided comments, which were incorporated prior to submission.

Candidate's signature _____

Supervisor or senior author's signature to confirm role as stated in (2) _____

DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women

Paula Dominguez-Salas[1,2], Sophie E. Moore[1,2], Darren Cole[3], Kerry-Ann da Costa[4], Sharon E. Cox[1], Roger A. Dyer[5], Anthony J. C. Fulford[1,2], Sheila M. Innis[5], Robert A. Waterland[6], Steven H. Zeisel[4], Andrew M. Prentice[1,2], Branwen J. Hennig[1]

Affiliations

[1]MRC International Nutrition Group, London School of Hygiene & Tropical Medicine, London, UK

[2]MRC Keneba, MRC Unit, The Gambia

[3]MRC Human Nutrition Research, Elsie Widdowson Laboratory, Cambridge, UK

[4]Department of Nutrition, Gillings School of Global Public Health and School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[5]Department of Pediatrics, University of British Columbia, Vancouver, British Columbia, Canada (RD, SMI)

[6]Departments of Pediatrics and Molecular & Human Genetics, Baylor College of Medicine, USDA/ARS Children's Nutrition Research Center, Houston, Texas, US

ABSTRACT

Background: Animal models show that periconceptional supplementation with folic acid, vitamin B12, choline and betaine can induce epigenetic differences in offspring phenotype mediated by changes in DNA methylation. In humans, altered DNA methylation patterns have been observed in offspring whose mothers were exposed to famine or who were conceived in the Gambian rainy season.

Objective: To understand seasonality of methylation patterns in rural Gambian women we studied natural variations in dietary intake of nutrients involved in methyl-donor pathways and their impact on the respective metabolic biomarkers.

Design: In 30 women of reproductive age (18-45 y), we monitored diets monthly for one year using 48h-weighed records to measure intake of choline, betaine, folate, methionine and vitamins B2, B6 and B12. Blood concentrations of these nutrients, as well as *S*-adenosylhomocysteine (SAM), *S*-adenosylmethionine (SAH), homocysteine and dimethylglycine (DMG) were also assessed monthly. Seasonality was modelled using truncated Fourier series.

Results: Dietary intake of vitamin B2, folate, choline and betaine varied significantly by season; with the most dramatic variation seen for betaine. All metabolic biomarkers showed significant seasonality, with the highest fluctuations in B6 and folate. Correlations between dietary intakes and blood biomarkers were found for B2, B6, active B12 (holotranscobalamin) and betaine. We observed a seasonal switch between the betaine and folate pathways and a limiting role of B2 in these processes.

Conclusions: Our data suggest that environmental influences impacting on maternal diet are measurable more reliably by plasma biomarker status than diet. This is likely to affect methyl-donor supply during pregnancy.

4.1.1. INTRODUCTION

The 'Developmental Origins of Health and Disease' hypothesis (DOHaD) is built on a substantial body of evidence that early life environment (especially nutritional) has lifelong effects on human health (109). The current challenge is to identify the underlying molecular mechanisms; epigenetic regulation is a strong candidate. Epigenetic modifications of DNA convey stable alterations in gene expression potential, that are not mediated by changes in DNA sequence. DNA methylation, involving the addition of methyl-groups ($-CH_3$) to cytosines within cytosine-phosphate-guanine (CpG) dinucleotides, appears to be the most stable epigenetic modification (110). Hence, DNA methylation patterns established during early ontogeny can persist for life (19).

One-carbon metabolic pathways supply methyl-groups for all biological methylation reactions (**Figure 4.1**). Methylated folate (5,10, methylene tetrathydrofolate) and betaine, a metabolite of choline, are the immediate substrates providing methyl-groups to re-methylate homocysteine and form methionine (111). Methionine is adenylated to *S*-adenosyl-methionine (SAM), which donates methyl-groups for DNA methylation, converting to *S*-adenosyl- homocysteine (SAH) in the process. SAH is then metabolized to homocysteine which can be re-methylated to methionine or enter the trans-sulfuration pathway leading to cysteine. Dietary supply of folate and choline, as well as of vitamins B2, B6 and B12, acting as enzyme cofactors are key to these processes.

Animal studies have demonstrated that periconceptional supplementation of the maternal diet with betaine, choline, folic acid and vitamin B12 can affect the establishment of DNA methylation patterns, altering phenotype and gene expression of the offspring (112). In humans such a causal association has not been described, however, periconceptional undernutrition during the 'Dutch Hunger Winter' was associated with decreased DNA methylation of the insulin-like growth factor (IGF2) gene in the offspring in adulthood (74). In The Gambia, a periconceptional micronutrient supplementation trial

showed association between periconceptional maternal nutrition and offspring DNA methylation, genome-wide and at imprinted loci (104, 105). Also, Gambian children conceived during the rainy season (August-September) had significantly greater methylation at five metastable epialleles (MEs) than those conceived in the dry season (79). MEs are genomic loci whose epigenotype is established stochastically in the early embryo and maintained thereafter across all tissues.

Figure 4.1: One-carbon metabolism

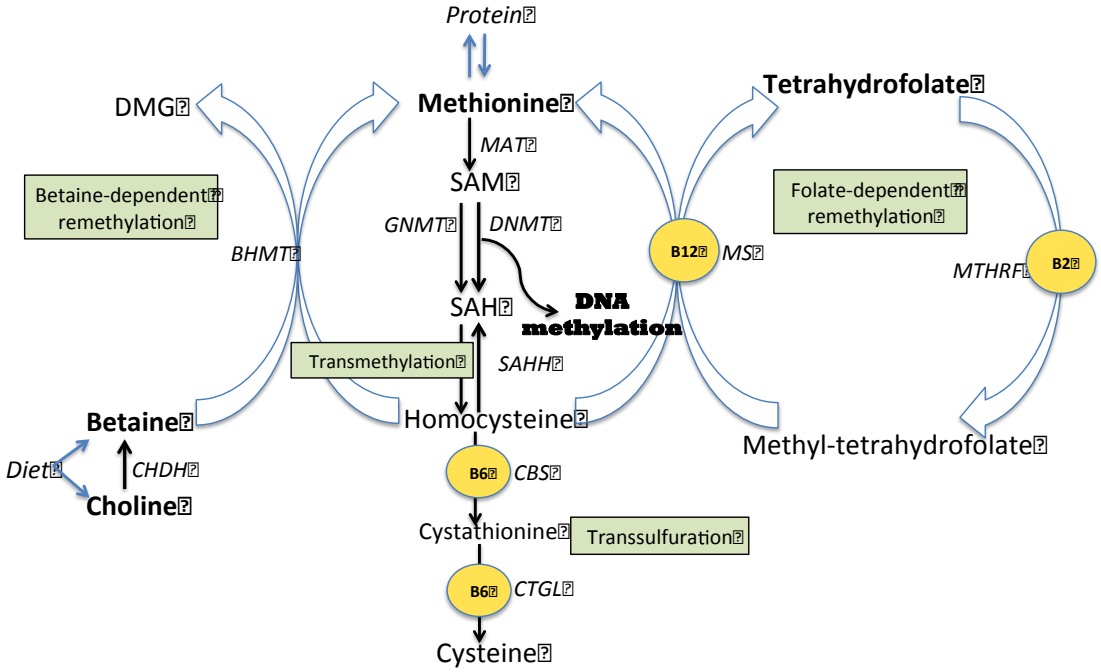


Figure modified from Dominguez-Salas *et al*. 2012 (111).

BHMT, betaine-homocysteine methyltransferase; CBS, Cystathionine-β-synthase; CHDH, choline dehydrogenase; CTGL, Cystathionine-γ-lyase; DMG, dimethylglycine; DNMT, DNA methyltransferases; GNMT, glycine N methyltransferase; MAT, methionine adenosyltrasferase; MTHFR, Methylenetetrahydrofolate reductase; MS, methionine synthase; SAH, S-adenosylhomocysteine; SAHH, SAH hydrolase; SAM, S-adenosylmethionine.

Methyl-groups released from conversion of SAM to SAH are also used for other methylation reactions (e.lg. proteins).

Seasonal variations in food supply and metabolic demand in rural Gambia provide a natural experiment to study effects of maternal dietary intake on epigenetic programming in humans. Here, we investigate the effects of seasonality on dietary intake and blood concentrations of methyl-donors, cofactors and related metabolites in rural Gambian women. We aimed to evaluate whether seasonal variation can explain the differences in DNA methylation status previously reported (79).

4.1.2. SUBJECTS AND METHODS

4.1.2.1. Study population

This observational study was conducted between July 2009 and June 2010. Non-pregnant women of reproductive age (18-45 years) from three villages in the rural area of West Kiang (Jiffarong, Janneh Kunda and Keneba), The Gambia, were invited to participate. The West Kiang Demographic Surveillance System, run through the MRC Keneba field station, was used as the sampling framework and the population sample (10 per village) was obtained by random selection, proportional to five age groups (18-45 years) to ensure all age groups were represented. Exclusion criteria included confirmed pregnancy, severe anaemia (<70 g/l), menopause, contraceptive use or planning to move away from the village during the course of the study. All blood samples were tested for haemoglobin concentration to determine anaemic status and malaria parasites, and participants were treated at MRC Keneba, when necessary. Any woman excluded during the course of the study, e.g. who became pregnant or chose to withdraw, was replaced by a new participant from the same age-group and village. Thirty women were visited each month throughout a single calendar year. Each visit involved two days of dietary assessment and collection of a fasted blood sample on the third day.

The Scientific Coordinating Committee of MRC Unit, The Gambia, granted scientific approval and the joint Gambian Government/MRC Ethics Committee (SCC/EC 1151) and

the London School of Hygiene and Tropical Medicine Ethics Committee (EC 5525) granted ethical permission for this study. Following community approval, informed written consent was obtained from each woman prior to participation.

4.1.2.2. Dietary intake measurement

Dietary records

The foods eaten in West Kiang are mostly grown locally. The diet typically consists of a staple (refined white rice, millet or maize) with a sauce made from a limited number of ingredients, such as vegetable/palm oil, groundnuts, green leaves, fish or vegetables (91). Full details of the local diet have been described elsewhere (88). Two main meals are eaten daily, one in the early afternoon and the second in the evening. Breakfast, often cereal porridge, left-overs from the previous day, or tea and bread in a more Westernised style, is eaten by some, but not all women. Food intake of each participant was determined by direct 24-h weighed dietary record on two consecutive days each month, using standardised procedures. On each day of assessment, women were visited in their household early in the morning by a trained fieldworker who then stayed with the subject throughout the day to weigh and record a description of all foods eaten, including meals and 'inter-meal' snacks. Since, families in this community eat from a common shared bowl, participants were supplied with a standard plastic container to separate and weigh their individual portion (scale Salter 1020 with Aquatronic Feature 5kg, 1 g accuracy). After taring the container, staples, sauces and any other foods were added one-by-one, with the container weighed after each addition, then weighed again to record the individual components of any left-over food at the end of the meal. Total food intake was thus calculated. Quantitative information on all recipes was obtained by weighing all individual ingredients, including water, as well as the cooked weight of the food (scale Salter

Brecknell WS 15 kg, 5 g accuracy). The food records and recipes also identified all of the different foods, including sauces, spices and condiments consumed.

Food composition

Contemporaneously, samples of 98 of the most commonly eaten foods in West Kiang were collected at different times of the year from the local markets (Appendix X (A,B)). Up to eight samples of each food were combined in a single composite sample. Food samples were prepared as consumed (raw or cooked), blended and freeze-dried (Edwards Modulyo EF4, USA) to constant weight before preparing composite samples for analysis of the target nutrients. Total choline (coefficient of inter-assay variation (CVi) = 4.7%) and betaine (CVi = 3.5%) were determined by liquid chromatography/electrospray ionization-isotope dilution mass spectrometry (LC/ESI-IDMS) (113), including the following choline compounds: free choline, glycerophosphocholine, phosphocholine, phosphatidylcholine and sphingomyelin, at UNC-CH Nutrition Obesity Research Center, University of North Carolina, USA. Vitamin B2 (CVi = 4.4%), B6 (CVi = 6.3%) and methionine (CVi = 4.3%) were determined by gradient high performance liquid chromatography with fluorescence detection against calibration standard solutions of known concentration (in-house methods), whilst folate (CVi = 7.6%) and vitamin B12 (CVi = 7.8%) were analysed by a surface plasmon resonance inhibition assay (Biacore Q assay) in a commercial laboratory (Eclipse Scientific Group Ltd., Cambridgeshire, UK).

Dietary intake calculation

The two sets of dietary data (dietary records and food composition) were entered into DINO ('Diet In Nutrients Out'). DINO is a dietary assessment software application specifically designed by MRC Human Nutrition Research (HNR), Cambridge, UK, to

calculate dietary intakes of Gambian foods. Intakes of energy, folate, B2, B6, B12, choline, betaine and methionine among the women in this study were thus calculated.

4.1.2.3. Blood biomarker measurement

Blood biomarker status measurements included folate, B2, B6, B12, active B12 (holotranscobalamin), choline, betaine and methionine, as well as total plasma homocysteine (tHcy), SAM, SAH and DMG. A fasted blood sample was taken in the morning of the day following the two-day dietary assessment. The samples were transported on ice to the MRC Keneba laboratory within an hour and immediately centrifuged, aliquoted and frozen at -70°C. Red blood cells (RBC) were washed three times with physiological saline solution (0.9% w/v NaCl), and stored at -40°C. Plasma tHcy (CVi= 1.7%), methionine (CVi = 1.1%), cysteine (CVi = 1.4%) (114), choline (CVi = 5.41%), betaine (CVi= 2.66%), DMG (CVi = 4.30%) (115), SAM (CVi = 1.0%) and SAH (CVi = 3.0%) were measured by liquid chromatography-tandem mass spectrometry (LC-MS-MS) in the Nutrition and Metabolism Laboratories of the Child and Family Research Institute at the University of British Columbia (CFRI/UBC), Canada, as previously reported (114, 115). Plasma vitamin B6 (CVi pyridoxal = 10.0%, CVi pyridoxal phosphate = 8.0%, CVi pyridoxic acid = 13.0%) was also measured by LC-MS-MS at CFRI/UBC based on the method of Midttum et al. (116), and plasma B12 (CVi = 3.3%), active B12 (CVi = 4.5%) and folate (CVi = 7.4%) were determined using a microparticle enzyme intrinsic factor assay and ion capture assay, respectively, with an AxSyM analyser (Abbot Laboratories, Chicago, IL), also at the CFRI/UBC laboratory. B2 status in RBC was assessed by the erythrocyte glutathione reductase activation coefficient assay (EGRAC), which measures the ratio of glutathione reductase activity in the presence and absence of added flavin adenine dinucleotide (CVi = 3.1%). EGRAC was performed on a microplate at MRC HNR (117). Higher EGRAC values denote greater B2 deficiency.

4.1.2.4. Statistical Analysis

Statistical analyses were performed using Stata 11.0 (StataCorp, College Station, TX). All dietary intake variables and blood biomarkers showed evidence of a positively skewed distribution and were therefore logarithmically transformed and are reported as geometric means. Biomarkers of nutrient status were considered as falling into two main groups for analysis: i) those that were consumed in the diet and also analysed in blood (i.e. folate, B2, B6, B12, choline, betaine and methionine, and ii) those which are here considered as 'functional biomarkers' of one-carbon metabolism, namely SAM, SAH, tHcy and DMG. Two additional variables were created, the SAM/SAH and the DMG/betaine ratio, to summarise the methylation activity and the betaine-pathway activity, respectively. The cut-off for statistical significance was considered at $P<0.05$.

To assess the reliability and stability of each dietary and biochemical (blood biomarker) variable, autocorrelations were calculated from multiple measurements of each woman one to five months apart. These were performed on 'detrended' values (i.e. adjusted for seasonality so that the observed difference between the measurements represented the sum of real variations in intake and/or biomarkers over time, as well as errors in measurement, and was not reflecting an underlying seasonality).

Trends along the year were calculated for diet and biomarkers separately. These were established using random effects multi-level models through Generalised Least Squares (GLS) regression to account for the repeat observations for each woman. The Fourier series approach was used to model patterns of seasonality (108). Two pairs of terms were judged to be sufficient for all variables, except for choline and betaine dietary intake and plasma B12, each of which required four pairs of terms. The EGRAC values reflect B2 status and are expressed as $1/x$ for seasonal trend graphic representation. To express the effect size of the seasonality, a coefficient of cyclic variation (CCV) was calculated, by square root of half the sum of the squared coefficients of the Fourier terms.

The differences in dietary intake and biomarker concentrations were also calculated comparing the peak of the rainy/hungry (Jul-Sept) and dry/harvest (Feb-Apr) season.

Finally, dietary predictors of biomarker concentrations were investigated cross-sectionally across all months, by pairwise matching dietary intake with the specific blood biomarker status by simple regression analysis. Multiple linear regression was then used to assess the relationship between dietary intake and the 'functional' biomarkers (which had no direct equivalent measured in the diet).

4.1.3. RESULTS

4.1.3.1. Study population

A total of 86 women were approached, of whom 62 consented and took part in the study. The mean age of the study participants was 31 years (range 18-45 years). Results for the dietary assessments of women who became pregnant during the course of the study were retained, but all biochemical analysis of blood samples taken after conception were removed from the data set as pregnancy can alter one-carbon metabolism (118). Therefore, there were complete dietary intake assessments for 28-30 women for each of the 12 months and fasted blood samples for 20 to 29 women per month. The average duration of participation in the study was 6 months.

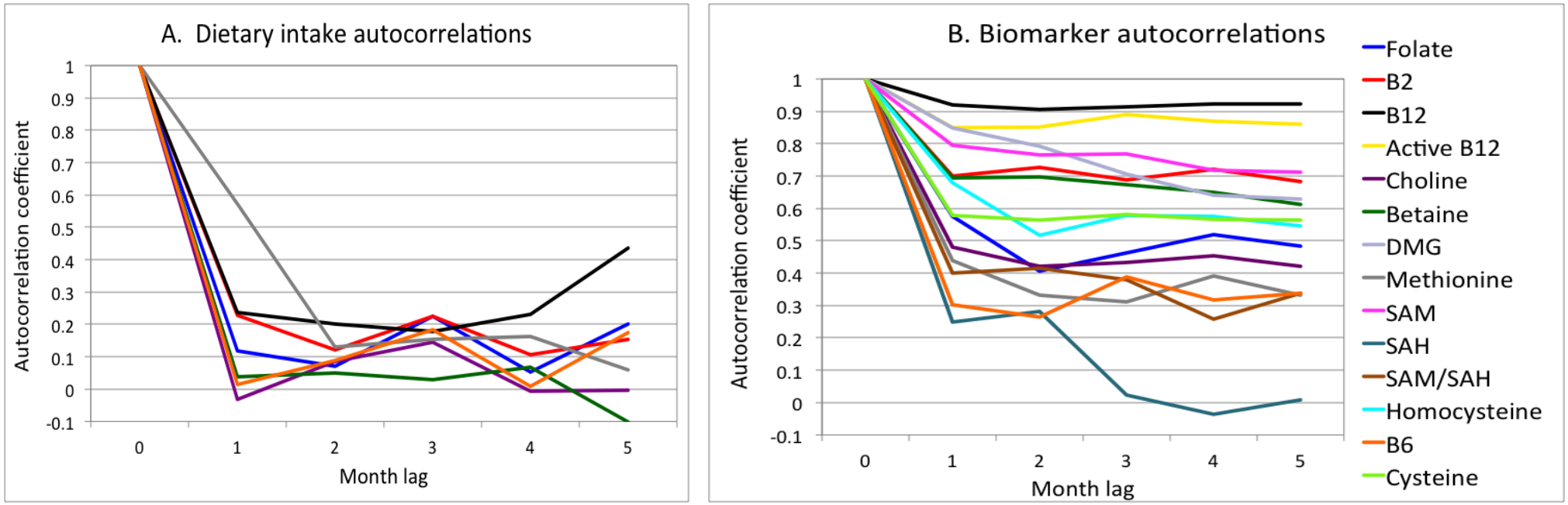
4.1.3.2. Dietary intake

Dietary data were collected for a total of 706 days, which included 1143 different recipes. The mean daily weight of food consumed (excluding water intake) was 1553 g (95% CI 1489-1619 g/d), with no significant variation over the year (p=0.78). The overall mean of energy intake was 1830 kcal/d (95% CI 1712-1816), with no significant variation over the year (p= 0.60).

The results of the autocorrelations to establish the reliability (precision of a single measurement or extrapolation back to zero lag) and stability (variation with time not related to seasonality) of the dietary intakes are shown in **Figure 4.2**. Additionally, repeatability of measurements on two consecutive days ranged between 40.1% and 60.1% (CCVs are shown in **Table 4.1)**. The dietary intakes of folate, B2, B6 and choline were significantly below the current international recommendations (see Table 4.1), expressed as Estimated Average Requirement (EAR). The EAR is the daily dietary intake level of a nutrient expected to satisfy the needs of 50% of a population group. The dietary intakes surpassing these recommendations are also in Table 4.1, and were below 30% for all except for B12 (100%) and methionine (98.9%).

The monthly variation in dietary intakes, expressed as percentage deviation from the overall geometric mean, are illustrated in **Figure 4.3, panels A, B, C and D**. The annual variation in the intakes of folate ($p=0.0038$), B2 ($p=0.016$), choline and betaine (both $p<0.0001$) was statistically significant. The results show that the intakes of B2 and folate decreased over the rainy season until September, followed by an increase, to reach their highest intake towards the end of the dry season peak (May). Choline and betaine intake was highest in September, with second peaks in December and April respectively. Methionine, B6 and B12 intake, conversely, did not vary significantly throughout the year ($p>0.05$). Among those nutrient intakes which showed significant variation over the year, substantial differences between both seasons were also shown for betaine, folate and B2 (Table 4.1) but not for total choline. B2 and folate were higher during the peak of the dry season (February-April, 12.1% and 14.9% respectively) and betaine was higher during the peak of the rainy season (July-September, 32.4%).

Figure 4.2: Reliability and stability of dietary intakes and blood biomarker concentrations for substances under study



Autocorrelation (internal correlation) of each variable measurement with measurements for the same woman one to five months apart (lag 1-5), adjusted by seasonality

n dietary intake measurements = average of 29 women x 12 months x 2 days; n biomarker measurements= 316 (293 for B2)

DMG, dimethylglycine; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine.

Table 4.1: Dietary intakes of one-carbon metabolites of Gambian women of reproductive age and international intake recommendations for women

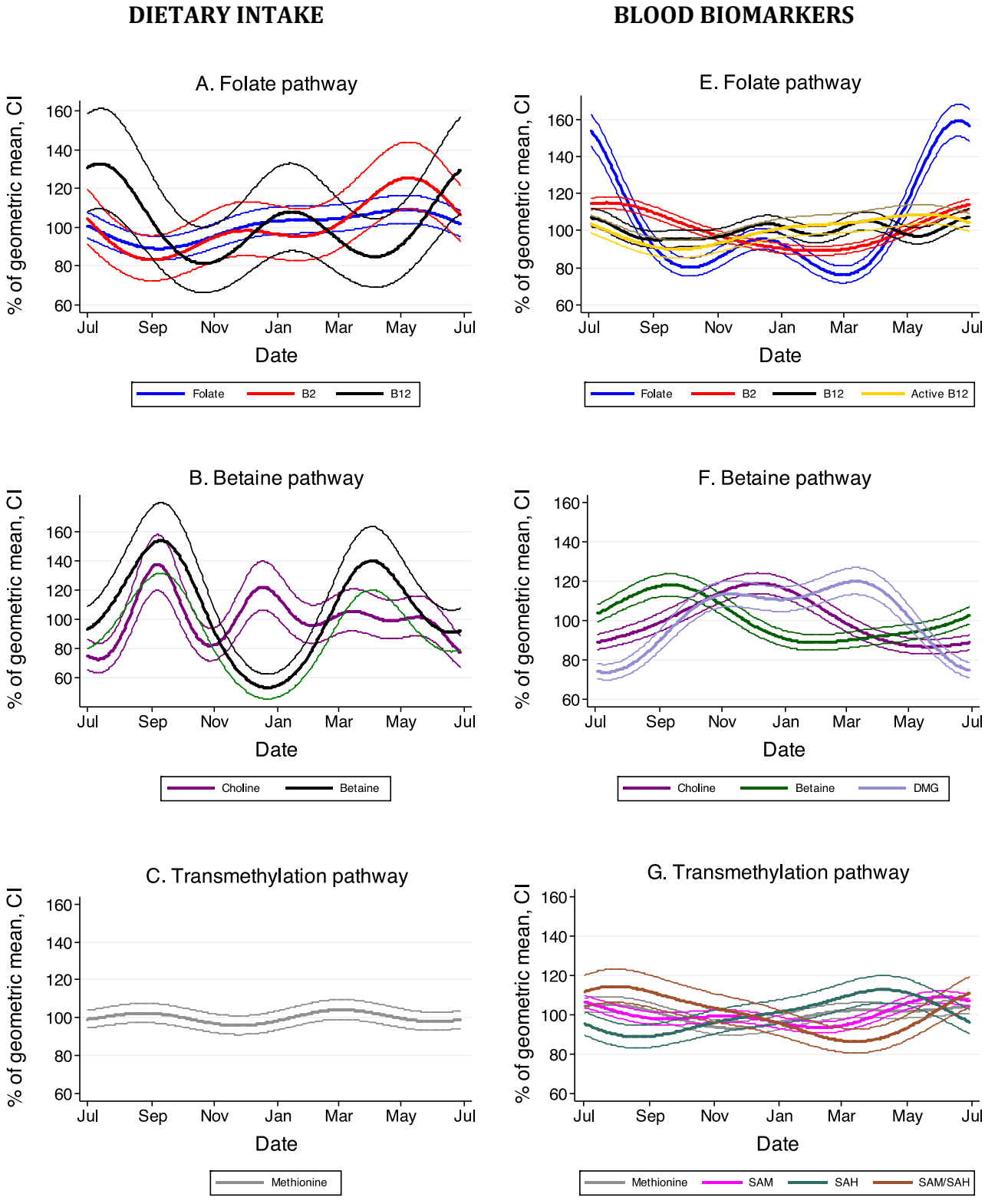
Nutrient [Units]	**12 month GM** (95% range)	**EAR** (prevalence of adequacy)	**GM Rainy season** (95% CI)	**GM Dry season** (95% CI)	**Difference between seasons** (%)	**Repeatability of measurement on two consecutive days** (%)	**CCV** (%) (CI 95%)
Folate [µg/d]	131.7 (69.2-251.5)	320 (1.2%)	120.5 (112.3-129.2)	138.4 (129.0-148.6)	-14.9*	52.1	6.6 (3.3-9.9)
B2 [mg/d]	0.30 (0.08-1.18)	0.9 (6.2%)	0.27 (0.23- 0.31)	0.30 (0.26-0.35)	-12.1	60.1	12.3 (5.4-19.2)
B12 [µg/d]	2.7 (0.4-21.0)	2 (100%)	2.8 (2.2-3.5)	2.6 (2.1-3.3)	5.8	44.9	14.7 (4.7-24.7)
Total choline [mg/d]	155.2[1] (60.6-396.7)	425[2] (2.8%)	158.4 (142.0-176.7)	158.4 (141.9-176.7)	0.1	52.9	15.1 (10.4-19.8)
Betaine [mg/d]	33.5 (6.9-163.3)	N/A	42.3 (35.3-50.6)	28.6 (23.9-34.2)	32.4*	48.8	31.8 (24.2-39.4)
B6 [mg/d]	0.92 (0.49-1.75)	1.1 (28.7%)	0.88 (0.81-0.95)	0.93 (0.86-1.01)	-6.4	40.1	4.3 (1.0-7.6)
Methionine [g/d]	2.03 (1.23-3.38)	0.83 (98.9%)	2.06 (1.94-2.18)	2.11 (1.99-2.24)	-2.7	44.7	2.4 (0.1-4.8)

n= average of 29 women x 12 months x 2 days; GM = Geometric mean; 95% range = GM ± 2SD; *p-value <0.05

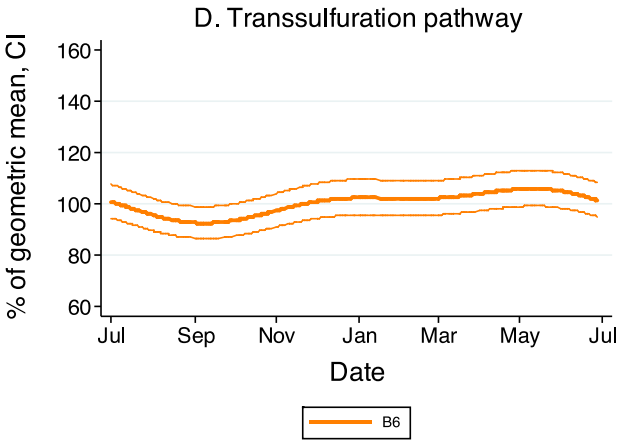
EAR: Estimated Average Requirement (119, 120): the daily dietary intake level of a nutrient expected to satisfy the needs of 50% of a population group; Repeatability = measurement of the reliability calculated on how similar the intakes between two consecutive 24-h weighed record measurements were in each subject; CCV= Coefficient of cyclic variation, calculated as the square root of half the sum of the squared coefficients of the Fourier terms; Rainy season 'Hungry'= July-September; Dry season 'Harvest'= February-April

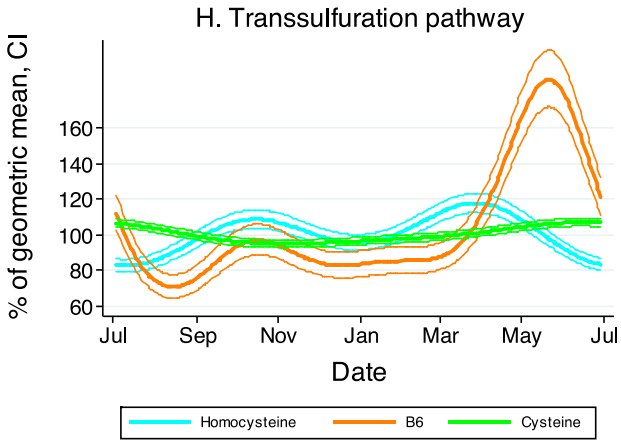
[1]52.6% Fcho, 15.6%Gpcho, 3.3% Pcho, 37.6%Ptcho, 0.8%Sm (Percentage contribution to total choline from free choline, glycerophosphocholine, Phosphocholine, Phosphatidylcholine, Sphingomyelin). [2] Adequate Intake (AI): dietary intake believed to be adequate for everyone in the demographic group to maintain health, established where no sufficient data to establish EAR are available, only an AI has been set for choline (21) [3] Recommendations are for Methionine + Cysteine are 19 mg/kg/d, calculated for an average weight of 55 kg

Figure 4.3: Seasonal trends of methyl-donors and cofactors as per dietary intake and blood



DIETARY INTAKE **BLOOD BIOMARKERS**





biomarker concentrations, expressed as percentage of geometric mean of all measurements (between July 2009 and June 2010)

n dietary intake measurements = average of 29 women x 12 months x 2 days; n biomarker measurements= 316 (293 for B2); 95% CI
B2 status is shown as the inverse of EGRAC assay results; DMG, dimethylglycine; Hcy, homocysteine; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine.
P-value for intake of folate = 0.0038; B2 = 0.016 , betaine and choline <0.0001; B6, B12 and methionine >0.05
P-value for biomarkers: B12= 0.03; all others $p<0.001$).

4.1.3.3. Blood biomarkers

During the 12 months of study, 316 plasma samples were obtained for analyses. Fewer RBC samples (a total of 293) were prepared for analysis of B2, as it was not always possible to obtain the second heparinised blood sample required. The autocorrelation calculations for biomarkers showed that, with the exception of SAH and DMG, all were stable (i.e. they correlate almost as well at 5 months as at 1 month lag) (Figure 4.2). Different degrees of reliability were apparent: the measurements of B12 and active B12 were highly reliable, whilst others, particularly SAH, showed poorer reliability (i.e. correlation was weak even at 1 month lag).

The geometric means and CCVs of biomarker concentrations for this group of Gambian women over the one-year study period are shown in **Table 4.2**. The annual variation in blood biomarker concentrations as a percentage variation from the geometric mean are illustrated in Figure 4.3, panels E, F, G and H. All blood biomarkers exhibited large, significant annual variation (B12 had a p-value =0.03 and all the other biomarkers $p<0.001$). However, the actual seasonal pattern of variation differed in timing and amplitude (magnitude of variation from the geometric mean, as indicated by the CCV), between the various biomarkers measured, with highest seasonal variation for plasma folate, B6 and the DMG:betaine ratio. Plasma folate and B6 status were highest in May-June, at the end of the dry season and lowest from September to March (Figure 4.3, panel E and H). B2 status declined progressively from the beginning of the rainy season peak (July), until January, then improved again. Betaine reached its peak later in the rainy season (September) (Figure 4.3, panel E). Plasma choline showed higher concentrations in November while B12 and methionine, in contrast to the other nutrients, remained relatively stable over the year (Figure 4.3 F, E and G, respectively). Regarding the functional biomarkers, the SAM:SAH ratio was highest in July, and lowest in March, with the seasonal fluctuation in the ratio explained largely by the seasonal change in SAH. Plasma tHcy showed an opposite change, with the lowest and highest plasma tHcy

concentrations found in July and April respectively. Plasma DMG concentration was highest between November and March, month at which DMG:betaine ratio reached the highest peak too.

As for dietary intake, Table 4.2 also shows the concentration differences between the peak of the rainy and dry season for each biomarker. Comparison of the biomarker concentrations according to these showed that the plasma SAM, betaine and folate concentration, and SAM:SAH ratio were higher , while the plasma SAH, tHcy, DMG, DMG:betaine ratio, B12 and B6 were lower in the rainy than in the dry season (Table 4.2). A lower EGRAC value during the rainy season denotes a less deficient status and thus higher B2 level. Cysteine, choline and methionine exhibited some variation between seasons but this represented less than a 1.5% difference between seasons. Effect of village was significant for intakes of B12 (p-value= 0.0023) and betaine (p-value= 0.0024), as well as for plasma concentrations of choline (p-value= 0.0039), folate (p-value= 0.0131), methionine (p-value= 0.0022) and SAM:SAH ratio (p-value= 0.0038). However the effect of village did not confound other effects.

4.1.3.4. Dietary intake and blood biomarkers

Next, the relationships between the dietary intakes of the nutrients under study and their blood concentration were regressed cross-sectionally to determine the predictive ability of dietary intake on blood biomarkers. Additional comparisons were made between selected nutrients to investigate interactions in specific pathways.

Positive associations between dietary intakes and respective plasma biomarkers were not observed for folate, choline (with total choline or any of its components) or methionine (data not shown). Significant positive associations were found between the dietary intake and blood biomarkers for B6 (β: 0.166, 95% CI: 0.017-0.316), betaine (β: 0.037, 95% CI: 0.006-0.069) and B2 (β: -0.038, 95% CI: -0.064- -0.0112, where a higher EGRAC indicates greater deficiency thus explaining the negative coefficient). Dietary B12

showed no significant association with the plasma total B12, but was positively associated with plasma active B12 (β: 0.043, 95% CI: 0.015-0.071). Plasma concentrations of folate were negatively associated with the combined intake of choline and betaine (β: -0.214, 95% CI: -0.297- -0.131). Plasma concentrations of choline and betaine were not, however, associated with folate intake, but were inversely associated with the intake of B2 (β: -0.065, 95% CI: -0.103- -0.028 and β: -0.058, 95% CI: -0.096–0.021).

To explore the potential impact of the dietary intake of the selected nutrients on one-carbon metabolism more generally, we also determined the strength of the associations between dietary intake and functional blood biomarker status. No statistically significant associations were found between the plasma SAM:SAH ratio or plasma SAM concentration and the intake of any of the nutrients under study. Plasma SAH showed association with B12 intake (β=0.039, 95% CI: 0.006- 0.072). Plasma tHcy and DMG concentrations were correlated with the combined intake of choline and betaine (β= 0.121, 95% CI: 0.060- 0.182 and β= 0.088, 95% CI: 0.004-0.172, respectively. As found for plasma choline and betaine, the plasma DMG was also inversely associated with the dietary intake of B2 (β=-0.086, 95% CI: -0.152- -0.020).

Table 4.2: Mean concentrations of blood biomarkers throughout the year and by season

Biomarker [Units]	**12 month GM** (95% range)	**Cut off for Adequacy** (prevalence adequacy)	**GM Rainy season 'Hungry'** (95% CI)	**GM Dry season 'Harvest'** (95% CI)	**Difference between seasons** (%)	**CCV** (%) (CI 95%)
Folate [nmol/l]	14.9 (6.6-33.4)	>10 (121) (82.3%)	13.4 (12.2-14.9)	12.5 (11.4-13.9)	6.7	23.8 (20.5-26.5)
B2[1] [EGRAC coefficient]	2.4 (1.5-3.7)	<1.3 (122) (0.7%)	2.3 (2.1-2.4)	2.6 (2.5-2.7)	-14.1[1]*	9.0 (7.6-10.4)
B12 [pmol/l]	342.3 (139.1-842.2)	>220 (123) (80.7%)	333.7 (285.9-389.6)	353.1 (302.5-412.1)	-5.8*	3.5 (1.9-5.1)
Active B12 [pmol/l]	70.9 (20.0-251.5)	>37 (124) (85.8%)	67.2 (55.9-80.8)	76.3 (63.5-91.7)	-13.5*	6.2 (3.7-8.8)
Choline [µmol/l]	8.2 (4.7-14.0)	>10 (120) (20.9%)	8.2 (7.6-8.8)	8.3 (7.7-8.9)	-1.3	11.1 (8.9-13.4)
Betaine [µmol/l]	40.4 (20.9- 78.4)	-	45.7 (41.7-50.1)	36.5 (33.3-39.9)	20.1*	9.8 (7.6-12.0)
DMG [µmol/l]	3.3 (1.1-9.3)	-	2.9 (2.5-3.4)	3.4 (2.9-4.0)	-18.4*	16.3 (13.4-19.3)
DMG:betaine	0.08 (0.03-0.26)	-	0.06 (0.05-0.07)	0.09 (0.08-0.11)	-49.0*	23.3 (20.6-25.9)
Methionine [µmol/l]	29.1 (19.6-43.3)	-	28.9 (27.3-30.7)	29.4 (27.7-31.1)	-1.5	3.9 (2.0-5.8)
SAM [nmol/l]	90.3 (59.1-137.8)	-	90.3 (85.3-95.5)	86.3 (81.6-91.2)	4.4*	4.6 (3.0-6.1)
SAH [nmol/l]	12.2 (6.8-21.9)	-	11.1 (10.5-11.9)	12.9 (12.1-13.7)	-15.9*	7.8 (4.7-10.9)
SAM:SAH	7.42 (3.8-14.5)	-	8.1 (7.6-8.7)	6.7 (6.2-7.2)	17.6*	9.4 (5.8-12.9)
Homocysteine [µmol/l]	8.03 (4.2-15.3)	<15 (125) (98.8%)	7.9 (7.2-8.7)	8.1 (7.4-9.0)	-2.9	10.2 (7.9-12.5)
B6 [nmol/l]	39.7 (14.3-110.6)	>20 (120) (91.1%)	33.8 (29.8-38.2)	36.0 (31.8-40.7)	-6.6	26.2 (22.4-30.0)
Cysteine [µmol/l]	224.9 (170.8-296.1)	-	217.1 (207.9-226.7)	219.0 (209.9-228.6)	-0.9	4.1 (3.0-5.2)

n= 316 (293 for B2); * p-value <0.05

GM = Geometric mean; 95% range = GM ± 2SD

Rainy season 'Hungry'= July-September; Dry season 'Harvest'= February-April

CCV= Coefficient of cyclic variation, calculated as the square root of half the sum of the squared coefficients of the Fourier terms

[1] For the B2 assay, higher values of EGRAC reflect higher B2 deficiency

(-) is indicated where no suitable cut-off could be identified

4.1.4. DISCUSSION

To our knowledge, this is the first study to comprehensively address dietary intakes and blood biomarkers relevant to one-carbon metabolism in an African population. Our results reveal that, as hypothesised, in this group of rural subsistence-farming non-pregnant Gambian women, dietary intakes and blood biomarker status of one-carbon-related nutrients and metabolites vary widely across the year. These seasonal fluctuations may provide the biological underpinnings of the seasonally-dependent changes in DNA methylation of metastable epialleles recently reported by us in this same population (79).

4.1.4.1. Dietary intake

Contrary to studies conducted several decades ago where energy intake was shown to be considerably lower during the rainy/hungry season (126), total food weight and energy intake in this study remained stable throughout the year. Several factors may contribute to this, including now-common financial remittances from family members living abroad that permit purchase of staple foods in the rainy/hungry season. High levels of energy expenditure in the farming season still perpetuate the historical rainy/hungry versus dry/harvest season divide, but it seems less severe nowadays.

Unlike overall energy intake, the qualitative composition of the diet varied over the year, resulting in significant monthly variation in micronutrient intakes. The dietary intakes of B2, folate, choline and betaine show clear seasonality, whilst methionine, B6 and B12 intakes remain stable over the year (Figure 4.3). A possible explanation is that the major dietary source of protein (fish), which was also a major source of dietary methionine and B12, remained relatively constant month by month, while more perishable crops varied more throughout the year. B6 is nearly ubiquitous in foods. An additional factor may be the month of Ramadan, which fell between August 20th and September 20th 2009 during our study period. This may have contributed to the peaks in

dietary intake of betaine and choline with foods such as bread, potatoes and eggs, which are consumed more frequently when breaking the fast (20). The dietary intakes of folate, B2, and choline were less than a half, and intake of B6 was more than three quarters of the current EARs.

Published data on seasonal variation in the intake of these nutrients in this population are available only for B2 and folate from 1979 to 1981(95). The deficient intake of B2 in both studies is consistent with the scarcity of milk and meat in the diet (88). B2 intake in the present study is lowest during the rainy season, which is consistent with dietary patterns found three decades ago (95). The seasonal variability in folate intake, however, is less consistent with the seasonality previously reported, perhaps explained by potential secular changes in the agriculture patterns (e.g. cultivation and consumption of green leaves in home gardens throughout the year). Oranges, which were identified as a source of folate in the past, particularly in the rainy season, were not consumed by any of the women in the study sample. However, these differences between the seasonal trends in both studies must be interpreted with caution. There might have been real differences due to genuine changes in the seasonal behavioural/dietary patterns but these might also just be methodological differences. In this respect, studies (and therefore food composition lab analysis assays) were 30 years apart, which could lead to technical differences in the measurements. Also this study was based on food analysis specific to the region under study (prone to a different types of error, as discussed in Section 4.4.4.), as opposed to using generally accepted food composition tables. These factors might lead to differences on either direction. Limited data are available on the dietary intake of these nutrients among women of reproductive age elsewhere in West Africa. Studies in urban Mali (127) and Burkina Faso (128) have reported similar low folate intakes, but higher intakes of B2 and B6, and lower B12 intakes, than in our population, possibly explained by differences in geographic location (e.g. river access and thus fish consumption), and proportion of purchased foods. The Adequate Intake (AI) level for choline in women is 425 mg/day (120); betaine recommendations are not yet fully

established. The intake values in this study were less than 40% of published AI. Whereas choline intake is known to be low in low-income countries (129), the 2005 NHANES survey found that adequate choline intake is often not achieved in the US either (130). Little is known about dietary betaine intakes, which in the present study had a geometric mean of 33.5 mg/d (95% range of 6.9 -163 mg/d). For reference, studies in the US have reported a range of betaine intake of 113-315 mg/d for adult women (131).

The repeatability of dietary intake between two consecutive days was good, suggesting that the women tended to eat similar foods in terms of dietary intake from one day to another, potentially from the same stocks. However, the stability from one month to another changed very quickly after adjustment for seasonality (Figure 4.2), indicating a low reliability in the measure of habitual intake. This could possibly be due to differences in cash availability at different times, leading to variations in food availability variation independent of season. Whatever the explanation, it is clear that the dietary intake trends should be interpreted with caution. This strongly favours the assessment of plasma biomarker concentrations over dietary intake.

4.1.4.2. Blood biomarkers

All plasma biomarkers measured showed significant seasonal trends throughout the year (Figure 4.3). The periods of peaks and troughs for the different biomarkers do not generally coincide. Experimental studies have shown a reciprocal interplay between the folate-dependent remethylation and betaine-dependent remethylation of homocysteine to methionine, such that deficiency of either folate or choline (and hence betaine) leads to increased utilization of the alternate pathway. (132). Notably, there were increases in DMG at a time when folate and B2 decreased – and vice versa – (**Figure 4.4)**, suggesting reciprocal switching from the folate-dependent to the betaine-dependent homocysteine remethylation pathway. High DMG may indicate remethylation of tHcy via betaine – homocysteine methyltransferase (BHMT) (Figure 4.1). Betaine is both a metabolite of

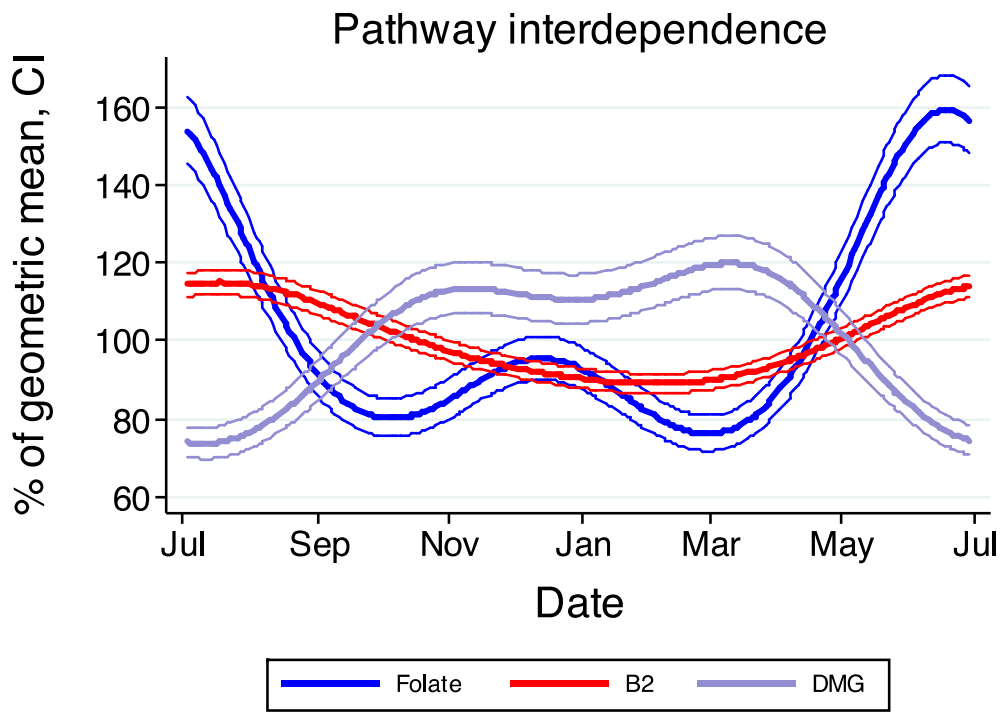
choline and also present in the diet, converting into DMG. The inverse relation between betaine and DMG gives good evidence of utilisation of betaine for remethylation and the DMG:betaine ratio also supports this (Figure 4.3, panel F). The status for B2 is highly deficient in this population. The seasonal nadirs in B2 status may therefore impair the folate pathway, leading to an activation of the betaine pathway. The effect B2 appears to play could be due to its role as cofactor for methylenetetrahydrofolate (MTHFR), which catalyses the formation of 5-methyltetrahydrofolate needed for homocysteine remethylation (Figure 4.1) (133). Whereas B12 could have a similar effect as cofactor for methionine synthase (Figure 4.1), in the folate pathway, blood concentrations of B12 do not seem to be limiting in this population. Furthermore, plasma tHcy decreases when B6 rises, suggesting a higher activity of the transsulfuration pathway. This pathway leads to cysteine, an important sulphur amino acid, and precursor of glutathione, via the vitamin B6 dependent enzyme cystathionase β-synthase (Figure 4.1).

Higher availability of methyl-groups and cofactors, as seen in increased levels of folate, betaine and B2, over the rainy season may explain why SAM is also higher at this time, leading to a higher SAM:SAH ratio, often viewed as a 'methylation index'. SAM converts into SAH by cleavage of methyl-groups during biological methylation reactions. The substantially higher methylation index (17.6%) would support a higher methylation capacity during the rainy season, consistent with previously reported DNA methylation at specific loci in Gambian children conceived during this season (79).

Bates *et al.* (95) described the widespread B2 deficiency in pregnant and lactating women in The Gambia 30 years ago as well as very similar annual trends. The poor status of B2 detected reflects the poor intakes observed, but the present study shows a much more marked deficiency than that reported in the previous studies, which were conducted between the 1970s and 1980s. The early EGRAC analyses were done using the original Glatzle method at the Dunn Nutrition Unit in Cambridge (no reference available) with some 'mid-stream' modifications based on later existing literature (134, 135). This method was later modified to avoid values of EGRAC<1, which appeared when samples

were saturated or nearly saturated (leading to reaction inhibition), by modifying the pre-incubation conditions, but also shifted the EGRAC towards higher values. There is evidence that changes in cofactor-induced reactivation conditions during the move to an automated laboratory method yields higher EGRAC values that consequently tend to overestimate the prevalence of B2 deficiency (122, 133).

Figure 4.4: Seasonal trends of folate, B2 and DMG blood concentration, expressed as percentage of geometric mean of all measurements (between July 2009 and June 2010)



Interdependence of folate-dependent and betaine-dependent pathways
n biomarker measurements= 316 (293 for B2); 95% CI
B2 status is shown as the inverse of EGRAC assay results; DMG, dimethylglycine
P-value blood B2, folate and DMG <0.0001

The stability and reliability of our measurement of biomarkers are generally better than those for the dietary intakes measurements. The comparatively reduced reliability of SAH and DMG measures could be due to either strong biological intra-individual fluctuations or to sample processing issues. SAM may convert to SAH during sample processing, and this conversion could be more strongly noticed in SAH than in SAM due to

the difference in total concentrations. However, every effort was made to avoid any delays in processing, thus we regard this to exert minimal effects on our measurements.

4.1.4.3. Dietary intake and blood biomarkers

The overall goals of this study were to examine: 1) whether dietary intakes predict nutritional blood biomarker status and contribute to the concentration of functional biomarkers, and; 2) how the seasonality observed in blood biomarkers reflected the seasonality observed in dietary intakes. Dietary intake was only directly associated with the blood biomarkers betaine, B2, B6, and active B12. The weakness of the associations between the monthly trends in biomarker concentrations and dietary intake may be explained by blood concentrations being subject to physiological utilisation (interactions between supply and demand) and complex regulation, or possibly by the inherent error associated with dietary assessment. As discussed above, biomarker concentrations can also be affected by compensatory mechanisms linking the folate- and betaine-dependent pathways, both methionine and homocysteine (which remethylates into methionine) playing pivotal roles. For example, blood concentrations of betaine are sustained by dietary intake and choline oxidation, which can in turn be obtained by diet or formed endogenously. Premenopausal women have the capacity to synthesise phosphatidylcholine via oestrogen-mediated induction of the phosphatidylethanolamine-N-methyltransferase (PEMT) gene (136). Conversely, use of betaine to provide methyl-groups takes place simultaneously and is related to the use of the folate-dependent methionine resynthesis cycle (137). As is apparent from the choline and betaine example, the metabolism of the different nutrients is highly interrelated, with deficiency or abundance of one substance affecting the metabolism of the others, probably to ensure maintenance of a constant supply of valuable methyl-groups. Furthermore, methyl-tetrahydro-folate requires its methyl-group being transferred for example from serine or glycine. Finally, homocysteine can be removed from the system by conversion to cysteine

via the transsulfuration pathway. SAM, the ultimate methyl-donor providing methyl-groups for DNA methylation, is also a positive allosteric regulator of cystathinine B synthase involved in transsulfuration.

4.1.4.4. Strengths of this study

This is the first comprehensive study of one-carbon metabolites in an African setting. Very detailed information was collected for close to 30 women every month for one year. In this population, the profound seasonal changes in diet quality combined with seasonal variations in metabolic demand due to very high levels of physical activity in the rainy/farming season create a special opportunity to examine the inter-relationships between dietary supply and circulating methylation-pathway metabolites in a manner that could not readily be achieved through experimental dietary manipulations. High levels of deficiency of several key micronutrients (vitamins B2, B6, folate and choline) also permitted us to study a situation in which the methyl-donor pathway may well be substrate-limited. Interestingly our data appear to show profound differences from studies of rural vegetarian women in India who are highly B12 deficient and relatively folate replete (138). Closer comparisons of these populations may provide additional insights into nutritional influences on DNA methylation.

4.1.4.5. Limitations of the study

As with most studies that attempt to assess human nutrient intakes in free-living conditions, there are several limitations. Although the weighed record method is viewed as the gold standard for quantitative dietary assessment (118), it is by no means perfect and behavioural changes from the daily practice could have taken place (139). Food composition data can also introduce error, due to limited number of samples and the high number of steps required in the processing. The random error inherent to any dietary

assessment method could have attenuated correlations (140), as clearly suggested by the reliability curves shown in Figure 4.2.

4.1.4.6. Conclusions

This study demonstrates that natural climate-dependent seasonality in food availability and dietary choices, especially in a micronutrient deplete population, creates an excellent 'experiment of nature' in which to examine linkages between diet, the metabolome and possible reproductive influences on DNA methylation in humans. This has allowed us to reach the following conclusions: 1) that, due to the biases (possible under-reporting) and imprecision inherent in estimating dietary intake, and due to the complex relationships between intake and nutrient availability for metabolic pathways, it is far preferable to assess plasma biomarker concentrations, as these provide a more accurate and more proximal indication of methyl-donor status; 2) that the previously described (137) reciprocal compensation between the folate-dependent and betaine-dependent re-methylation pathways may be driven by a moderate-to-severe vitamin B2 deficiency that is paradoxically worst in the dry/harvest season; 3) that, contrary to our initial assumption that the rainy/hungry season would put women at greatest risk of a deficient methyl-donor supply, the reverse was clearly apparent with lower homocysteine and higher methylation indices (SAM:SAH ratio); and 4) that these findings are consistent with our previously reported findings of elevated DNA methylation at metastable epialleles in individuals conceived during the rainy season (79).

ACKNOWLEDGEMENTS

We would thank the women of West Kiang who patiently participated in the study and acknowledge the enthusiastic work of the fieldworkers, lab technicians and nurses who collected the data and samples, and to all the staff at MRC Keneba in general. Also thanks to Dr. Dorothea Nitsch (LSHTM) and Dr. Chris Bates, Lorna Cox, Dr. Gail Goldberg

and Dr Celia J. Prynne, (MRC Human Nutrition Research, Cambridge) for their valuable advice and input to the study, and to Tongwen Wang (UNC) and Jannette King (UBC) for their assistance with laboratory analyses.

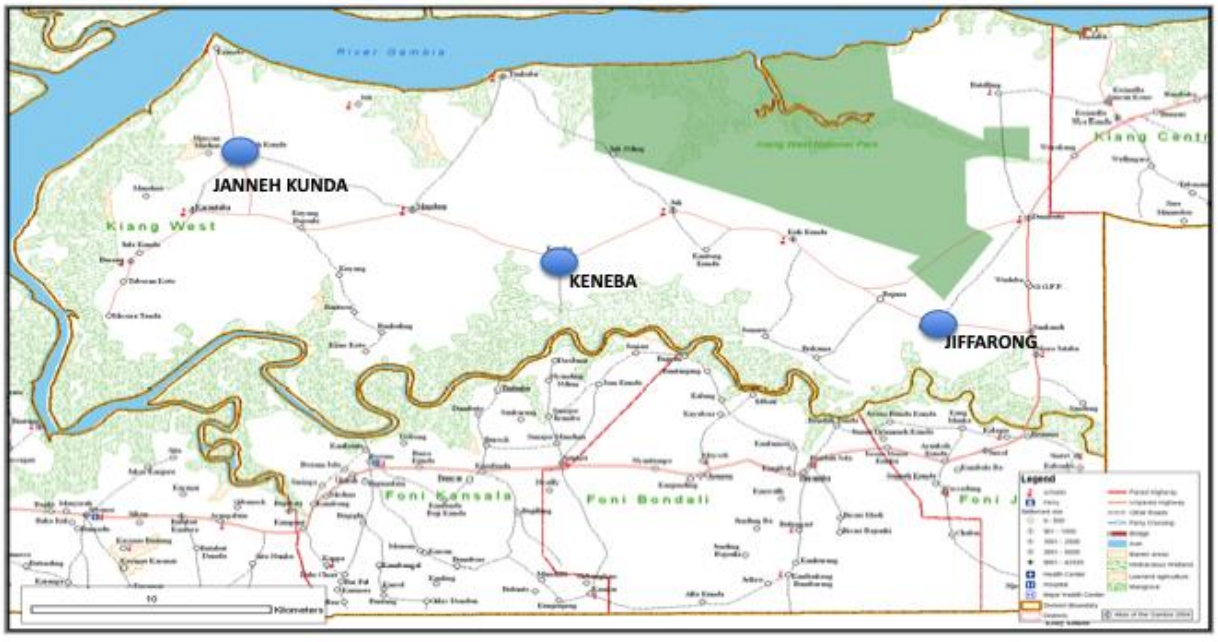
The authors' responsibilities were as follows: PDS, SEM, SEC, AJCF, AMP & BJH conceived and designed the study; PDS, SEM & BJH were responsible for conducting the research; PDS was involved in all hands-on experiments, and conducted the sample and data collection; PDS & AJCF performed the statistical analyses; DC gave technical support for dietary intake data processing; KAC, RAD, SMI & SHZ conducted the biochemical analyses; PDS drafted the paper and all authors critically revised and approved the manuscript. No conflict of interests associated was declared.

4.2.ADDITIONAL METHODS

4.2.1. Study population

The selection of the three villages from which participants were selected (Jiffarong, Keneba and Janneh Kunda) was based on several criteria. The aim was to capture as much variation as possible within the district. These villages are located at Eastern, Central and Western points within the region, as shown in **Figure 4.5**. The three villages selected differed according to location to main access routes (South Bank road), proximity to river and proximity to primary healthcare centres, although all villages are predominantly Mandinka.

Figure 4.5: Map of geographical location of the indicator group villages within West Kiang



Indicator group villages

Meetings were held with the authority ('alkalo') of each of the three villages to seek permission for conducting this study. The burden of the study group was high for the participants and for their families (i.e. two full-days of fieldworker presence at their compound, in addition to blood sample being collected monthly). Therefore, details of study and the protocol were carefully explained to the women and also to their families. Additionally, sessions of sensitisation were conducted to respond to the concerns raised regarding the frequency of the bleeding and the amount of blood.

A short LMP questionnaire, similar to that for the main group, was conducted. Women who did not see their menses for two consecutive visits had a urine sample collected for pregnancy testing (Quickvue One-step hCG Combo © dipstick test, BioMérieux SA, Marcy l'Étoile, France). If the result was positive, they were invited to MRC Keneba for an antenatal care visit and excluded from the indicator group if the pregnancy was confirmed. The gestational age was determined by ultrasound scan, to determine whether samples should be used. Participants from the indicator group were excluded from the main group recruitment, but if they became pregnant and fulfilled the eligibility criteria of conception timeframes, they were invited to enrol into the main group.

4.2.2. Dietary intake

<u>24-hour weighed records</u>

Field assistants were trained to conduct a very detailed assessment of dietary intake (two day 24-hour weighed intake records recipe collection), as described in Section 4.1.2.2 (Appendix III (B)). Each day, the field assistant would arrive at the participant's compound prior to the cooking/breakfast time and remain there throughout the day, or followed women to the field. For the recipe preparation (e.g. porridges, sauces or full meals such as "nyankatango" – rice normally cooked with pounded groundnuts and flaked fish), each ingredient was weighed separately (after washing, peeling and eliminating inedible parts), including added water, and all ingredients were recorded in the 'recipe

form' (Appendix VI), with information on the food sources (i.e. home garden, bush, shop, rural market, coast or other). The form included details on the potential for cooking smoke exposure (i.e. type of kitchen –ventilation).

At the beginning of the study, subjects were supplied with a standard plastic container that they kept for the length of the study, and used for the separation of their individual portion for measurement and so that their individual intake could be estimated accurately (as opposed to their normal practice of family shared bowls) (139). Components of the meal were added one-by-one progressively (generally staples were added first, followed by sauces, and then any other extra foods (pieces of vegetables, fish or meat)) and recorded separately in the dietary intake form (Appendix VII). Thus, each new measure included all the previously added foods, in addition to the container weight and the newly added food. Values for each individual food could then be calculated by difference with the previous measurement (taring method). If the portion was not finished, the inverse procedure was followed for the left-over weighing. Meat or fish bones were weighed first and subtracted from the initial amount of fish/meat. All other individual ingredients left-over were then measured and lastly the staple was weighed, indicating if it was still mixed with sauce or alone. If the woman expressed her will to continue eating her food afterwards, left-overs were again measured at that stage (recorded as "second attempt" in the dietary intake form – Appendix VII). Any snack or drink consumed during the day (other than water) was also weighed.

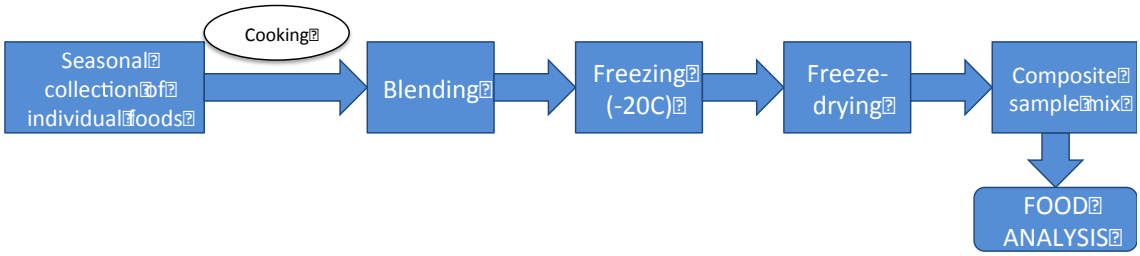
Food composition analysis

In parallel to the dietary fieldwork, a list with all ingredients used by the participants was compiled, with details of the most frequent source also recorded. The list was updated every month throughout the year. On average five samples (range 2-8) for each ingredient were collected (subject to availability), by the field assistants or by myself, for content analysis of methyl-donors and cofactors. Samples came mainly from the three villages of study, regional rural markets were also visited and some foods were obtained

from the coastal urban area. Particular attention was given to potential seasonal differences: two sets of samples, one during the rainy season and one during the dry season, were collected for those foods deemed to potentially experience composition changes across seasons, such as meat or rice. The decision was made depending mainly on whether the sources varied along the year (e.g. local *vs.* imported rice) or if they were literature reports of potentially important changes of composition (e.g. meat composition). The details on the collection and origin of each sample were compiled in the food collection form (Appendix XI, Form G02A).

The information on the processing was recorded on the laboratory processing form (Appendix XI, Form G02B). A flow diagram of the procedure for food preparation prior to analysis of methyl-donor and cofactor content is shown in **Figure 4.6**. Final 'composite' samples (n=98) were created to analyse and obtain an average composition value, by mixing together individual freeze-dried samples of the same ingredient but with different origins. The composite sample mix was prepared by combining equal amounts of the original foods (i.e. prior to freeze-drying). The amount of water needed to rehydrate the composite samples was calculated and sent to the analysing laboratories. At the University of North Carolina the powdered sample were analysed (analysis of choline compounds and betaine), whilst Eclipse Scientific Group rehydratated the samples prior to analysis (analysis of B2, B6, B12, folate and methionine). The main details on those analyses are given in Section 4.1.2.2. Total choline analysis included five different compounds: free choline (Cho), glycerophosphocholine (GPC), phosphocholine (Pcho), phosphatidylcholine (PTC), and sphingomyelin (SM). Betaine is also a choline metabolite but values are not included in the calculation of total choline since the conversion of choline to betaine is irreversible (141). The lower detection limits for quantification of substances under study were as follows: folate 0.021 μg/g, B2 0.5 mg/kg, B6 0.1mg/kg, B12 0.002 μg/g and methionine 0.02g/100g. Detection limits for choline compounds ranged between 1 and 40, with betaine detection being the most sensitive and GPCho detection being the least sensitive (113).

Figure 4.6: Food processing



Flow diagram of the processing steps prior to analysis of foods (ingredients)

Dietary intake calculation

Data collected in the recipe and dietary intake forms were entered into the DINO programme (version 2.1). This programme is an Access-based application, with several sections. The section of 'Food tables' includes food composition tables, namely those that already exist for some nutrients and the expansion based on the information from this project for methyl-donors and cofactors. It also contains the 'Nutrient Profile Calculator' for recipe calculation, based on the system used for the UK National Diet and Nutrition Survey. The section on 'Dietary Coding' allows for data entry using the taring system (for the meals) or portion sizes (for the snacks). Finally the 'Data Analysis' section provides result tables in Excel, including daily intakes, average portion weights and main contributors. Examples of data entry screens in DINO are shown in Appendix XII.

From the beginning of the dietary data collection (July 2009), the forms were entered and a list of all the food items coded was drawn up, for the collection of foods to be analysed. This list was extended as new records were entered. Recipe and dietary intake data were cleaned on data entry, as DINO does not allow for any missing value, or mismatching value (for example weight of rice + sauce being below the weight of rice only), or unusually high portion sizes. This 'check system' helped early detection of problems in the data collection that could be addressed by conferring with the field assistants conducting the work. This was the first time that DINO was used in real time for a project. While piloting the software, some issues were raised in order to improve its use, and also to better respond to the specific needs of the current study. These were

addressed through changes in the programme, to make it easier to use by data clerks and better reflect the intakes on an individual basis. For example, the 'Nutrient Profile Calculator' application was designed specifically for this study.

4.2.3. Blood biomarkers

Details of blood collection and sample analysis are given in section 4.1.2.3. The 10ml blood sample was collected into a 7.5 ml EDTA Sarstedt Monovette© and a 2.4 ml LithHep Sarstedt Monovette©.

Samples of whole blood were used to prepare malaria slides and haemoglobin concentration was measured with a Delphi portable haemoglobin meter (Delphi Industries, Auckland, New Zealand). Then, samples were spun at 2750 g and 4°C in a Hettich EBA 21 centrifuge for 10 minutes to separate plasma for aliquoting and later biomarker analysis. RBCs were washed 3 times in the LithHep tube with physiological saline solution (0.9% w/v NaCl), and then spun at 2750g for 15 minutes, prior to storage at -40°C. All sample collection processing methods followed protocols, as agreed with collaborating institutes and in line with local procedures at MRC Keneba.

B2 status was the measured as EGRAC on microplates. This method developed in – house at the MRC HNT is equivalent to the method on the Cobas FARA centrifugal analyser previously used at HNR, and based on the initial procedure of Vuilleumier et al. (117) for a Cobas-Bio centrifugal analyser. Measured EGRAC results were verified by internal cross-comparisons with the FARA method, but the microplate assay has better precision (unpublished data). The reagents and preincubation conditions used for both methods (microplate and Cobas-FARA) are the same, with the exception of the addition of a low concentration of detergent (Triton X-100) in the current method, which is needed to equalise the light-path in the microplate wells. For both methods, the activity of erythrocyte glutathione reductase is measured before and after stimulation by *in vitro* saturation with flavin adenine dinucleotide (FAD). The main difference between methods

is the absorbance reader and in this case, the reduction in absorbance at 340 nm was read with an iEMS temperature-controlled microplate spectrophotometer.

4.2.4. Data analysis

Dietary intake

During the data entry into DINO, a number of problems were encountered. For example, it is not unusual that the remaining rice and sauce from evening meals are mixed overnight and stored for breakfast the following morning. In such cases, the real proportions could not be calculated but a rice-sauce average proportion, from other data available of rice with a similar sauce for the same participant and at similar time of the year, was used instead.

Once the nutritional composition results from the freeze-dried food samples were available, these were entered into DINO (as composition per 100g of food) and data cleaning / quality control checks performed. On occasions, possibly due to an under-recording of the added water during the recipe observation (the water was usually added at various stages), an 'overconcentration' of the recipe (i.e. more water evaporated than actually added) was observed. To check the adequacy of the recipes concentration, the energy content of each recipe was plotted against the energy content information of the same 'standard' recipe available in DINO. For the energy calculation of the newly collected recipes, data available from The Gambia (96) and Mali (142) were mainly used, and from British foods otherwise (143). Outliers affected by overconcentration were replaced by the same recipe for the same woman at similar time of the year. When that was not possible, the recipes of other participant in the same village for the same month were considered. The software calculated automatically the daily intake of the substances under study.

In addition to the analysis included in the 'Dietary intake' section of the research paper II (Section 4.1.3.2), the correlations between pairs of the different dietary intakes of

methyl-donors and cofactors were calculated, after standardisation for total energy intake. Additionally, the principal contributors to the total population intake were calculated, by grouping the main intakes by ingredient or by type of recipe (e.g. different types of fish, rice preparations or green leave sauces).

Blood biomarkers

Full details of the analysis of blood biomarkers are given in the Methods section of the research paper II (Section 4.1.3.3).

Dietary intake and blood biomarkers

Full details of the analysis of dietary intake and blood biomarkers are given in the Methods section of the research paper II (Section 4.1.3.4)

4.3.ADDITIONAL RESULTS

4.3.1.Study population

To help ensure a constant sample size of 30 women per month, a replacement strategy was put in place, where women who withdrew or who were excluded were replaced by a new recruit from the same village and age group. The reasons for replacement were:

a) Pregnancy (n=14): pregnant women were removed from the study after confirmation of status by ultrasound. There are no known preferential dietary customs or major taboos (i.e. foods prohibited) specific to pregnancy in this population that could alter dietary intake qualitatively (126). A recent study in a comparable rural community in Burkina Faso has also shown no significant difference in the dietary intake between pregnant and non-pregnant women (144). Although it is acknowledged that intake could be reduced during early pregnancy in some women feeling sick, the dietary intake data of pregnant women were included

in the analysis until the moment where they were replaced. However, one-carbon metabolism is known to be substantially affected by pregnancy (145), due to hormonal and physiological changes and enhanced requirements, and therefore blood samples collected during the first months of pregnancy were excluded from the analysis. Thus, blood samples collected during pregnancy, as confirmed retrospectively using gestational age (due to the time delay of pregnancy detection), were excluded from the analysis.

b) Self-withdrawal (n=11): a higher withdrawal rate was observed than expected, seemingly as a consequence of the monthly bleeding (quantity and frequency). Out of the 11 women who withdrew, more than half (n=6) were from Janneh Kunda, whilst 3 were from Jiffarong and 2 from Keneba.

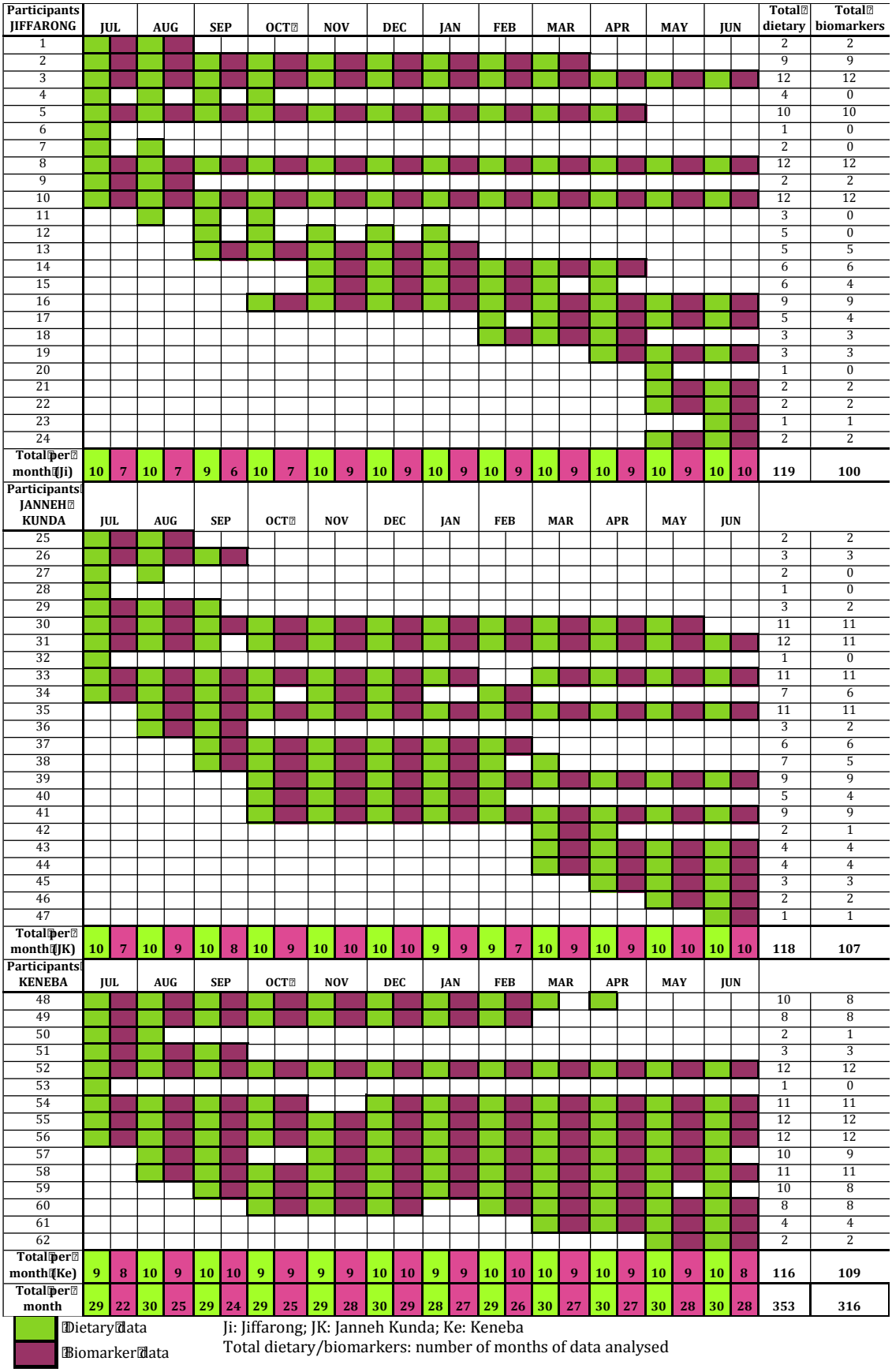
c) Moved away or long term travel (n=6): participants who could not be visited for more than one consecutive month were excluded from the study and replaced.

d) Menopause (n=5): these cases were identified during the first visit, which explains the smaller sample size for the first month. After this, special care was given to early menopause (or underestimation of date of birth, in a population traditionally without birth registry) and to ensure the participants were still seeing their menses when joining the study. As was the case for pregnant women (and for contraceptive users, see below), the information collected on dietary intake information was included, but blood biomarker data was disregarded in the analysis.

e) Known contraceptive users (n=1).

A small number of participants (n=4) refused to have their blood collected after their dietary intake on a few occasions. Together with the samples withdrawals, this led to a reduced sample numbers available for the biomarker analysis compared to dietary intake assessments. The details on follow-up with respect to both dietary intake and biomarker data points for each participant by month and village are shown in **Figure 4.7**.

Figure 4.7: Duration of follow-up of participants in the indicator group and sample size of dietary and biomarker data



4.3.2. Dietary intake

The effect of village of residence and participant age (in years) was assessed as covariates with respect to the dietary intake of each individual nutrient by multiple regression (GLS). An effect of village was seen for the intakes of B12 (Keneba = 4.08 μg/d, Jiffarong = 2.55 μg/d, and Janneh Kunda=2.08 μg/d; p-value = 0.0023) and betaine (Keneba = 40.27 mg/d, Janneh Kunda = 33.83 mg/d and Jiffarong = 28.48 mg/d, p-value = 0.0024), with intakes for both highest in Keneba. Age had no effect in any of the individual dietary intakes.

The monthly variation in dietary intakes, expressed as percentage deviation of the geometric mean is presented in the research paper II in Section 4.1.3.3. In addition, **Figure 4.8** shows the seasonal variation for the different choline compounds measured. These all contribute to the calculation of total choline but they were also looked at individually for a better understanding of their individual contribution to total choline intakes. All five compounds had significant annual variation. The graph shows that the choline peak in September is likely to be mainly driven by phosphatidylcholine and phosphocholine, whilst the peak in December seems to be driven by sphingomyelin and free choline, and the later smaller peaks, by free choline and phosphatidylcholine. The best sources of free choline were foods of vegetable origin (green leaves, potatoes and groundnuts), whilst for glycerophosphocholine, phosphocholine, phosphatidylcholine and sphingomyelin, main sources were of animal origin (egg, shrimp, milk or meat). The specific information on food composition is shown in Appendix X (B).

As discussed in research paper 1 (Section 2.2) the supply of methyl-groups for remethylation of homocysteine in one-carbon metabolism has two complementary pathways. Some of the different dietary methyl-donors and cofactors are needed in combination (i.e. folate, B2 and B12). Therefore, independently of seasonality and repeat measurements, the correlations between different intakes within the Gambian diet were assessed. The correlation coefficients, adjusted for energy intake, are shown in **Table 4.3**.

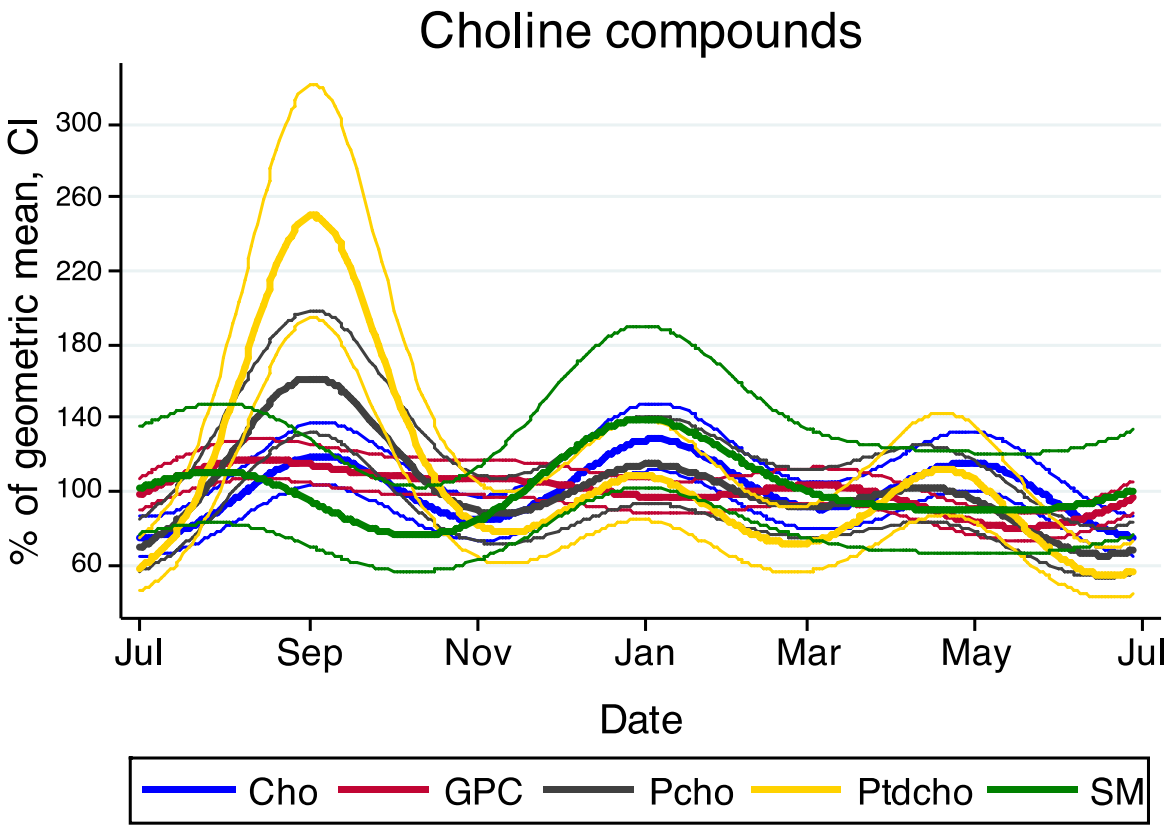
B12 is not correlated with any other nutrient whilst the other B-vitamins, B2, B6 and folate, seem to be significantly correlated, with coefficients ranging between 41.0% and 62.8%. The highest correlation was observed between choline and betaine. These correlations can be explained to a large extent by common sources of each nutrient contributing to the total intake. Major contributors (grouped by food type) to the dietary intake are summarised in **Table 4.4**.

Table 4.3: Correlation coefficients between methyl-donors and cofactors in the diet of rural Gambian women

Nutrient	B2	B6	Folate	B12	Choline	Betaine
B2	1					
B6	0.5132*	1				
Folate	0.4100*	0.6275*	1			
B12	0.0764	0.0145	-0.0142	1		
Choline	0.2471*	0.2961*	0.2832*	0.0989	1	
Betaine	0.1678*	0.1010	0.1504*	0.0054	0.3682*	1
Methionine	0.0903	0.2331*	0.2185*	0.0092	-0.0774	-0.0130

* p-value <0.05

Figure 4.8: Seasonal trends for dietary intake of choline compounds, expressed as percentage deviation of the median (July 2009 to June 2010)



Thick line: percentage of the geometric mean; Thin lines: 95% CI
Choline compounds: Free choline (Cho), glycerophosphocholine (GPC), phosphocholine (PCho), phosphatidylcholine (Ptdcho), and sphingomyelin (SM).
P-value intake Cho = 0.043; GPC <0.0001; Pcho = 0.005; Ptdcho <0.0001; SM = 0.036

Table 4.4: Main contributors to the methyl-donor and cofactor intakes in the diet of rural Gambian women

Nutrient	B2	B6	Folate	B12	Methionine	Choline	Betaine
1st contributor*	Rice	Rice	Rice	Fish	Rice	Rice	Leaf sauces
2nd contributor*	Milk	Millet	Millet	Meat	Fish	Groundnut	Bread
3rd contributor*	Millet	Leaf sauces	Mangoes	Milk	Millet	Beans	Fish

Ingredients were grouped by type of food (e.g. green leaf sauce includes all the sauces made out of leaves such as morango, kucha, jambo or others)
* 1st, 2nd and 3rd contributors to population intakes were calculated quantitatively

Wood was used as cooking fuel by all women observed. They generally used the same type of kitchen throughout the year, independent of the season: 80.5% of women cooked in closed kitchen of mud brick, which were poorly ventilated with just one door and without any window; 3.7% cooked in kitchens with some windows; and 7.9% cooked outside in the courtyard. The remaining 7.9% of women showed a variation in cooking place across the year (i.e. closed kitchen used more often during the rainy season and open air cooking done more often during the dry season).

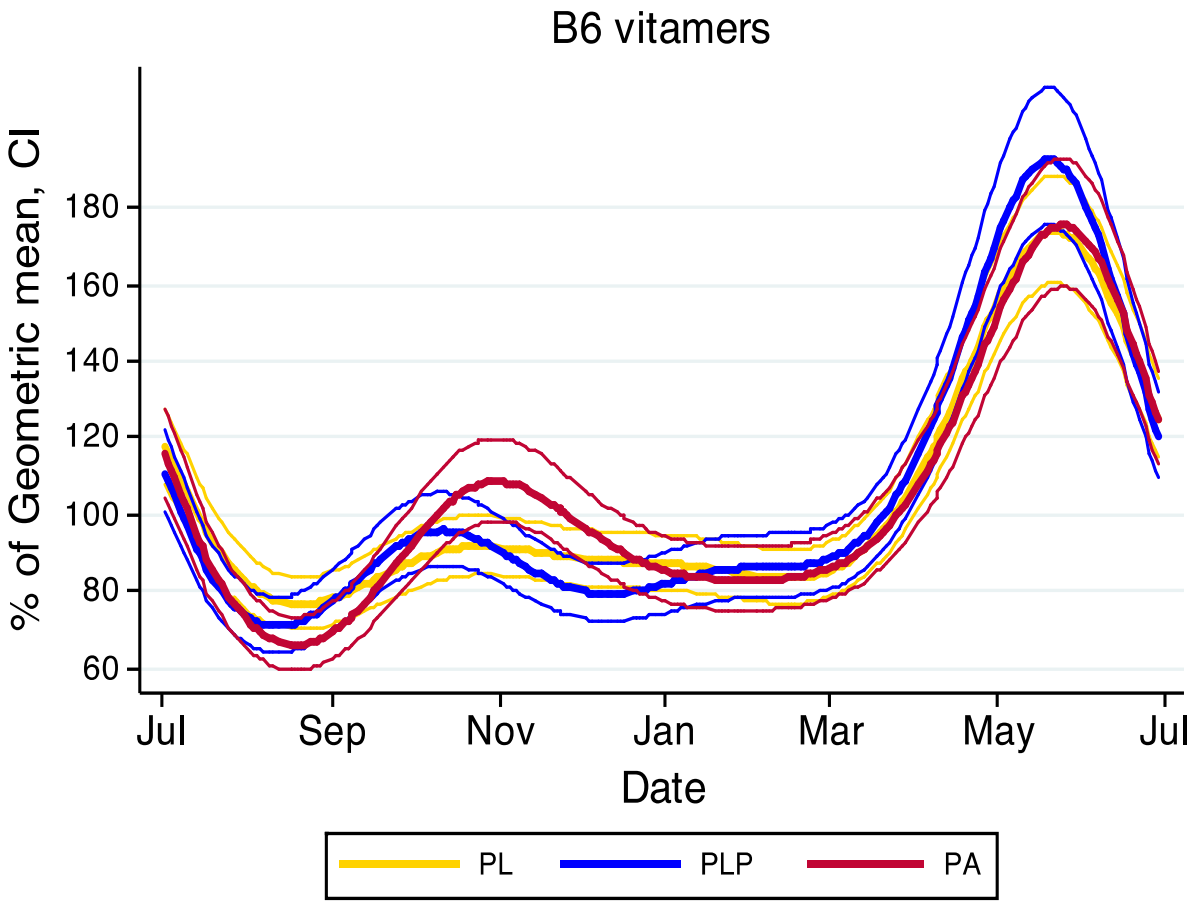
4.3.3. Blood biomarkers

The effect of village of residence and participant age (in years) was also assessed as covariates with regards to the blood biomarker levels. An effect of village was found for folate which was highest in Janneh Kunda (Janneh Kunda = 15.80 nmol/l, Keneba = 15.14 nmol/l and Jiffarong = 12.49 nmol/l; p-value = 0.032), and for SAM:SAH ratio, which was higher in Keneba than the other two villages (Keneba = 7.89, Janneh Kunda = 7.39, Jiffarong = 6.88, p-value< 0.0001) and lowest in Jiffarong in both cases. Choline (ß= 0.009 μmol/ per year, p-value= 0.001), betaine (ß= 0.010 μmol/l per year, p-value= 0.03), SAM:SAH ratio (ß= 0.007 per year, p-value= 0.001) and SAM (ß= 0.006 nmol/l per year, p-value= 0.029) levels seemed to increase with age, whilst methionine (ß= -0.005 μmol/ per year, p-value= 0.038) decreased. However, village and age were not included in the final analysis models because the effects observed made little difference to the overall results (details not presented).

As described in Section 4.1.2.3, participants were asked to fast overnight prior to blood sample collection. On 24 occasions (out of 316; 7.6%), women reported to have eaten. However, without the information on what was eaten, how much or when, it was decided not to take into account the fasted / non-fasted status in the overall analysis models.

Seven different vitamers exist for B6 (116). Three of those vitamers have been analysed for this study, namely pyridoxal (PL), pyridoxal-5'-phosphate (PLP) and 4-pyridoxic acid (PA). **Figure 4.9** shows the seasonal variation for the different B6 vitamers measured. There is a strong similarity in their variation throughout the year, although with different concentrations. The geometric mean concentrations were 3.76 nmol/l (95% CI 3.54-4.01 nmol/l) for PL, 27.68 for PLP (95% CI 25.04-30.59 nmol/l) and 7.77 (95% CI 7.07-8.53 nmol/l) for PA.

Figure 4.9: Seasonal trends for B6 vitamers in blood expressed as percentage deviation of the median (July 2009 to June 2010)



PL: Pyridoxal; PLP: Pyridoxal-5'-phosphate; PA: 4-Pyridoxic acid; all p-values <0.0001

The association of the different individual methyl-donors and cofactors with the main functional biomarkers was assessed in **Table 4.5.** SAM is a positive allosteric

modulator of cystathionine β synthase (CBS) (146), which means that SAM regulates the flux through the transsulfuration pathway which coverts homocysteine into cysteine. Hence cysteine was considered as an 'outcome' blood biomarker in this study, even if we did not investigate the dietary intakes. Table 4.5 shows that plasma B6 did not seem to predict any of the selected functional biomarkers (p-value>0.05), other than cysteine. Homocysteine was associated with all methyl-donors and cofactors, but the folate-pathway and betaine were associated with decreased tHcy, whilst choline and methionine were associated with an increase of tHcy, SAM:SAH ratio showed a similar pattern, although neither folate nor active B12, seemed to play a role. The DMG concentration was positively associated with choline and betaine concentrations, but also negatively with folate and B2.

Table 4.5: Regression coefficient of the individual blood methyl-donor and cofactors with the main functional biomarkers in plasma.

Methyl-donor/cofactor	**Functional biomarker**	**β coefficient**	**P-value**
Folate [nmol/l]	SAM:SAH ratio	0.100	0.060
	SAM	0.077	0.015*
	SAH	-0.027	0.628
	Homocysteine	-0.182	<0.0001*
	DMG	-0.119	0.019*
	Cysteine	0.051	0.019*
B2 [1/EGRAC]	SAM:SAH ratio	0.254	0.011*
	SAM	0.087	0.187
	SAH	-0.114	0.251
	Homocysteine	-0.219	0.018*
	DMG	-0.532	<0.0001*
	Cysteine	0.028	0.532
Active B12 [pmol/l]	SAM:SAH ratio	0.046	0.216
	SAM	-0.003	0.924
	SAH	0.2867	0.001*

Methyl-donor/cofactor	Functional biomarker	β coefficient	P-value
	Homocysteine	-0.198	<0.0001*
	DMG	0.020	0.715
	Cysteine	0.016	0.413
Choline [μmol/l]	SAM:SAH ratio	-0.331	<0.0001*
	SAM	0.051	0.301
	SAH	0.287	0.001*
	Homocysteine	0.213	0.002*
	DMG	0.257	0.001*
	Cysteine	0.088	0.008*
Betaine [μmol/l]	SAM:SAH ratio	0.164	0.015*
	SAM	0.027	0.559
	SAH	-0.139	0.042*
	Homocysteine	-0.293	<0.0001*
	DMG	0.430	<0.0001*
	Cysteine	-0.075	0.014*
Methionine [μmol/l]	SAM:SAH ratio	-0.2123	0.034*
	SAM	-0.039	0.445
	SAH	0.101	0.281
	Homocysteine	0.266	<0.0001*
	DMG	0.013	0.873
	Cysteine	0.112	0.001*
B6 [nmol/l]	SAM:SAH ratio	0.016	0.720
	SAM	0.021	0.398
	SAH	-0.007	0.871
	Homocysteine	0.060	0.079
	DMG	0.065	0.091
	Cysteine	0.048	0.004*

* p-value <0.05; Generalised Least Squares regression between log-biomarkers, adjusted by seasonality (sin1-cos2)

4.3.4. Dietary intake and blood biomarkers

Details of the results of relationships between dietary intakes and blood biomarkers are covered in the Results section of the research paper II (Section 4.1.3.4).

4.4. ADDITIONAL DISCUSSION

In this section, additional issues beyond the scope of the research paper II (Section 4.1.4) are discussed.

4.4.1. Study population

As discussed in Section 4.3.1, it was not possible to follow a stable population across the full 12-month period of study. This could have introduced random bias, potentially confounding the individual effect with the season effect, in those individuals who only participated for a specific month or season. This would have averaged out within a bigger sample but the indicator group sample size was limited. 'Subject' (participant ID) was fixed in the model to account for the repeated measures. This yields a conservative estimate of the season effect, which underestimates the effect, detecting only genuinely significant associations. This might have affected the ability to detect significant seasonality in some of the dietary intakes. All plasma biomarkers were shown to be seasonal, but the obtained estimates could be potentially less accurate (due to the reduction in power). A further drawback was the relatively small sample size, affecting the power for the detection of associations between the dietary intake and plasma biomarkers.

4.4.2. Dietary intake

In this section, further consideration of the factors affecting the dietary intake is given. With the exception of methionine and B12 (and betaine, which does not have a recommended intake), the dietary intakes observed in this population are considerably below the current recommendations. Few women reached the EARs for B2, choline and folate (6.2%, 2.8% and 1.2% respectively were beyond the EAR, see Table 4.1 in research paper II, Section 4.1.3.2). In comparison with other populations, mean intake of choline of women in other LIC such as Jamaica (278 mg/d) (129) was below recommended Adequate Intake of 450-550 mg/day (147) and similar to the choline intake of the lowest quartile of women in the USA (<293 mg/d in 1984 - Nurse's Health Study (148)). The mean intake in The Gambia was nearly half this value (155 mg/day). Mean intakes B2, B6 and folate of women in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort throughout 10 Western European countries were of 1.6 mg/day, ~1.2–2.3 mg/day and ~175–331 μg/day respectively in women (149), compared to 0.3 mg/day, 0.92 mg/day and 155 μg/day in The Gambia.

A statistically significant seasonal variation in dietary intake was shown for B2, choline, betaine and folate. B2 (mainly present in milk and meat) intake is lowest during the peak of the rainy season and increases from September onwards. This is not an isolated observation, as discussed in Section 4.1.4.2, but is consistent with previous findings (95). Based on personal observation, a potential explanation for the seasonality patterns is that it seems preferable not to milk / slaughter cows at the end of the dry season and the beginning of the rainy season, due to the poor condition the animals are at after a long period of scarcity of grass. The B2 intake is poor throughout the year. However, it is not unusual to see villagers sending containers of milk to the coast (personal observation). Thus, it is possible that milk could be sold at the coast in exchange of cash, reducing the local availability even during times of plentiful supply. For folate, the increase along the second half of the dry season seems to coincide with the brief local

availability of mangoes, which is consistent with previous findings (95). The post-harvest dry season coincides with the time when choline seems to be driven by free choline (Figure 4.8), which is more frequent in ingredients of vegetal origin. Conversely, the choline peak in September appears to be driven by the choline metabolites which are more present in foods of animal origin. One also has to be bear in mind that, as discussed in Section 4.1.3.2, Ramadan in 2009 fell in the middle of the rainy season (between August 20th and September 20th 2009) and might therefore have introduced bias. It has been reported by Gambian women that during Ramadan there is increased consumption of certain foods such as beans, vegetable oil dishes, fish, potatoes or meat (150). From observation of the annual trends (Figure 4.3) the choline and betaine peaks in September, could very well be driven by the Ramadan dietary practices. Conversely, methionine intake is very stable throughout the year and, B2, folate and B6 daily intakes seem mostly unaffected by the fasting.

The usual diet in rural Gambia relies on a relative small number of ingredients. This leads to a degree of overlap in the main sources of the different nutrients, concentrated in fish, meat, green leaves, groundnuts and milk (see food composition table in S1 in Section 4.1.3.2, Table 4.3 and Annex X). However some of the richest foods (eggs, meat, milk) are consumed very seldom by women, as they are expensive and only temporarily available. Furthermore, most of the nutrient-rich foods are not consumed in large quantities, resulting in less rich foods (i.e. rice or millet) being the major contributors to the diet. This probably explains the high intake correlations of B6, folate and methionine.

An effect of village was not considered in the regression model used to analyse dietary intakes (GLS regression) because of the little impact it seemed to have, but B12 and betaine, intake appeared to be higher in Keneba, followed by Jiffarong and finally Janneh Kunda. In terms of access to fish, which is the main B12 source, all three villages have access to river fishing catches to a certain extent. However, Jiffarong and Keneba have better access than Janneh Kunda to goods brought from the coast (e.g. a lorry which

regularly brings sea fish from the coast) and from the big rural markets on the main road (where a variety of dry fish is sold). Differences in cash availability could also play a role, with generally higher income for Keneba residents, due to the presence of the MRC as the main employer in the area, as well as the International Trypanotolerance Centre (ITC), also located in this village. The betaine sources that explain differences in betaine intake are bread (48% and 69% more was attributable to bread in Keneba than in Janneh Kunda and Jiffarong, respectively), probably due to higher bread availability in Keneba – there is a baker who makes bread almost every day- and cash availability to purchase it) and green leaves (27% and 61% more was attributable in Janneh Kunda than in Keneba and Jiffarong respectively). The higher consumption of green leaves in Janneh Kunda could be to a good organisation of communal gardens, together with more difficult access to markets. Therefore, they might be more likely to eat what they produce, as opposed to selling it.

In terms of cooking smoke exposure, the anecdotal data collected would not suggest much intra-individual variation (across season) but greater inter-subject variation instead. These observations could be biased due to the variable length of follow-up per participant. Owing to drop-outs, some participants were followed only during one season, and therefore there is insufficient longitudinal data across the year to assess smoke exposure accurately.

Given the limited range in dietary practices (ingredients and food preparation diversity) in the region, two days of dietary assessment were deemed to be sufficient (151). Such assessment is subject, as in most settings, to methodological limitations inherent to nutritional studies. Although 24-hour weighed records are considered the gold standard, this method is by no means exempt of problems, particularly behavioural changes. In a society where shared bowls is common practice, individual measurements become more challenging. The separation of the subject's portion, particularly when assessed by someone from their community, might have put the subjects under pressure to deliberately under-eat. Alternatively, being observed could lead to subjects over-eating as they do not have to share the food with others. The observation of the cooking can also

induce substantial behavioural changes such as simplification of the cooking procedures or, conversely, the preparation of a more complex (potentially better) meal. In either case, each person would have tendency to one or the other, and therefore this should have a constant effect along the year, with high inter-subject variation but much less intra-individual variation.

One way to assess obvious under- or over-reporting is to look at the calculated energy intakes among the women. The average reported energy intake was 1830 kcal/d. Previously, a total daily energy expenditure of approximately 1650 kcal/d has been reported in non-pregnant, non-lactating women (89, 152). This was measured by 24-hour whole-body calorimetry, with 60 minutes of compulsory exercise. The current measurements of intake would allow an additional 180 kcal/d of discretionary activity. This is inadequate during the intense physical activity of farming season (153), but is probably close to the requirement for the year as a whole. Therefore, although underestimation of food intake is a well-know limitation of dietary intake methods (139, 153), it is unlikely that the current estimates of intake represent an error of more than 10% of the average.

Although this is not thought to have had a major effect, food composition data could also introduce error in the dietary intake estimation, due to the limited number of samples and the high processing which was required (i.e. cooking, freeze-drying, preparing composite samples).

4.4.3. Blood biomarkers

In terms of secular trends, EGRAC coefficients are much higher (indicating greater deficiency) now than in the past (95). Although the poor biomarker status detected reflects the poor intakes observed, there is evidence that the move to an automated laboratory method yields higher EGRAC values that consequently tend to overestimate the prevalence of B2 deficiency (122, 133) compared to historic data available. In addition, it

has been argued that an improved B2 status, as measured by EGRAC, could take place during the rainy season when tissue resorption (weight loss) takes place, with potential release of B2 from tissue turnover (e.g. muscle) and transfer to enzymes such as glutathione reductase. Conversely, during the dry season, tissue restoration would lead to increased B2 demands, and would lower the circulating B2 (95).

The only previous study looking at plasma folate, B12 and total homocysteine, 12 years prior to the current study, was limited to a three month period during the dry season of 1997 (97). The geometric mean for folate (17.5 nmol/l) seems higher than in the current study (14.9 nmol/l). However, the previous study took place from April to June and in the current data, the monthly mean plasma folate concentration ranged between 13.91 in April and 23.32 mean nmol/l in June, suggesting that a lot of variability seems to exist across this time period, so these differences in values seem plausible. The same study reported substantially lower levels for B12 (262 pmol/l *vs.* 342 pmol/l) for both pregnant and non-pregnant women. This could potentially be due different analytical/sample collection and processing techniques used (IMx kit assay compared to AxSYM in this study). However, substantial equivalence between assays has been claimed by Abbot, manufacturer of both assays (154). If methodological reasons are ruled out, higher plasma concentration found in this sample population could be due to a higher B12 intake nowadays, as a consequence of increasing cash availability and/or a more regular supply of fish from the coast as mentioned above. However, no data are available on past B12 intakes and therefore it is difficult to ascertain what the reasons could be. In any case, and supported also by the dietary intake data, B12 levels are adequate throughout the year and are probably not limiting for one-carbon metabolism. The past values for tHcy were slightly higher (8.92 μg/l) than the values observed in the current study (8.03 μg/l), and this could be explained by the timeframes investigated and the variation over that period (mean homocysteine ranged between 9.88 μg/l in April and 6.94 μg/l in June).

Large 'phase' differences (i.e. with different patterns of peaks and troughs along the year) can be observed between the different biomarkers. This can be to a certain

extent explained by seasonal variability in the diet, but also by physiological use and turnover. This makes defining a globally 'advantageous' or 'restricted' time period of the year difficult. Therefore, the use of SAM:SAH ratio as a methylation capacity index to summarise one-carbon metabolism is very useful. In Table 4.5, it can be observed that the SAM:SAH ratio seems to be only positively associated with B2 and betaine concentrations, whilst negatively with choline and methionine concentrations. The opposite direction in the association between SAM:SAH and the different methyl-donors (e.g. choline and betaine), is difficult to explain. This finding is consistent with the correlations of choline and betaine for other biomarkers (i.e. SAH, homocysteine and cysteine) with the exception of DMG, for which both show positive association. In this respect, choline shows similar effects to methionine, suggesting that the explanation might be in the alternative biological functions (e.g. protein building) that these two metabolites have, thus making the interpretation more complex.

The strong negative association of B2 with DMG supports the observation based on the seasonal graphs (Figure 4.4) indicating that low B2 could lead to enhanced activity of the betaine-dependent pathway. B12 in this population is not limiting (i.e. beyond the established cut-off, Table 4.2 Section 4.1.3.3.) which might explain why it does not show an effect on the SAM:SAH ratio or in the shift to the betaine-pathway (increased DMG). However, B12 seems to be associated with SAH and homocysteine. Plasma B6 not did predict any of the selected biomarkers other than cysteine. Even if, from one-carbon metabolism, cysteine could be considered an 'outcome' of the transsulfuration (via B6 regulation), cysteine, as methionine, also comes from the diet and is present in body protein pools for other functions. Nevertheless, B12 concentrations seemed to respond to folate, choline, betaine, methionine and B6 concentrations.

Choline in plasma was measured as free choline. It has been recently observed that elevated total plasma choline (which is primarily phosphatidylcholine and sphingomyelin since they tend to occur as much bigger pool than free choline) was associated with lower

risk of NTD, not plasma free choline (155). The relationship between different choline pools is not fully understood.

4.4.4. Dietary intake and blood biomarkers

It seems somehow paradoxical that while some of the dietary intakes reported for Gambian women participating in this study are considerably below the current EARs, these are not accompanied by seriously deficient biochemical or physiological indices for most of the metabolites. An exception is intracellular B2. The EGRAC coefficient is very high (the average value in this group is almost twice the established cut-off), seemingly reflecting the very poor intake in this population (with an average intake three times below the EAR). Furthermore, there is a seeming lack of response of blood concentrations to dietary intakes for some of the nutrients, as shown by the weak or lack of correlation between values in intake and in blood. These mismatches between low dietary intake and yet comparably limited deficiency in blood biomarker level, with lack of correlation between both measurements could be explained by:

1) Measurement error inherent to dietary intake measurement. The limitations of the dietary assessment methodology have already been highlighted in Sections 4.1.4.2 and 4.1.4.5. Figure 4.2. clearly illustrated the poor reliability and limited stability existing for all dietary intake data collected, in contrast with the reliability and stability observed for those nutrients in blood. The local dietary practices (i.e. eating from a shared bowl) are undoubtedly disrupted by the procedures followed to measure individual intakes of the women in the study. Similarly, regular eating choices can be affected during the days of measurement by the presence of the field worker in the compound (e.g. participants eating more, or less, or preparing foods differently to usual practices in front of guests; or conversely oversimplifying the cooking on the measurement days to avoid the burden of the measurement) or could be

misrecorded if women decided not to report certain things eaten when not in front of the fieldworker. Beyond the actual recording of intake, the transformation of food intake into nutrients intake also can lead to error, from the calculation of nutrients during the food preparation to the variety in food composition. The study tried to overcome this by analysing local foods, prepared the local way, but this was not exempt to error either, due to the limited sample size used, and the high processing and weighting of the samples (when prepared, cooked, freeze-dried, analysed, etc.) that could introduce error at any stage. These limitations in the dietary measurements have consistently been observed in dietary assessment (156). With the clear potential for over/under reporting, the lack of good correlation between nutrient intake and blood level variables is unsurprising, and the choice of blood levels as better and more reliable status markers becomes undeniable; also

2) Biomarkers being subject to physiological utilisation and to tight regulation, which could not be measured based on the sampling method employed here. The dietary intake may not always directly reflect in blood for biological examples. An example of this are the blood concentrations of betaine and choline which can, as discussed in the research paper II in Section 4.1.4.2, also be explained by endogenous synthesis and turnover. Indeed, choline is not essential in women of reproductive age, as it can be synthetised from phosphatidylethanolamine via PEMT enzymatic action. This endogenously produced choline, as well as dietary choline, can transform into betaine. Yet, betaine intake is associated with blood biomarker concentration in this study, although this association is not very strong. However, this example illustrates that these substances are not solely dependent on diet, and given the interdependence of all the substances in the one-carbon metabolism, this could also have an effect on the levels in blood of other substances.

B2 as measured by the EGRAC assay is commonly reported to correlate well with dietary intake and also shows an association with dietary B2 in this study. The EGRAC assay (which alters in accordance with erythrocyte half-life), reflects availability of B2 stored intracellularly, giving a good estimation of long-term B2 status (118). Thus, while the other nutrients under study are assumed to be more directly related to recent intake, the B2 level is assumed to reflect longer-term intakes. However, the association observed is not unexpected since B2 can rapidly cross the RBC membrane. Folate, B12 and methionine blood levels seem unaffected by diet. As discussed above (Section 4.1.4.3), the metabolism of the different nutrients is highly interrelated, and the deficient levels of most could be due to the metabolism of others. For instance, methionine is an essential amino acid, which can be obtained from the body protein pool or through the remethylation of tHcy. Finally, genetic variation could modify folate and choline metabolism resulting in individuals whose nutrient turnover may differ (136). Such genetic variation needs further study in this population.

The associations observed for the functional biomarkers are also somewhat paradoxical. For example, it was surprising not to find correlations between SAM, SAH or tHcy with the dietary intakes of folate, B6 or B2 (157), especially given the apparent severe deficiency of B2 in this population. Only tHcy was found to be associated with the choline and betaine intake, but this association was positive. This is contrary to what has been previously reported in the Netherlands and in a US population prior to folate fortification (131, 158, 159). Thus the findings of this study should be considered with caution and need further assessment, e.g. if supplementation studies are to be considered in this population.

One of the main objectives of this study group was to provide dietary predictors of biomarkers that could be extrapolated to the main group, to estimate their intakes for a given blood concentration, in the next Chapter. The reliability and stability of some of the intakes and biomarkers needs to be considered (see Sections 4.1.4.1 and 4.1.4.2). Further,

the poor correlation between dietary intake and biomarker levels might make it difficult to extrapolate the data from the indicator group for use and estimation of the potential dietary intake in the main group. Dietary data was shown clearly to be more imprecise and variable than the biomarker measurements, which look more robust than the dietary intake. Therefore, the use of biomarkers in the main study seems more appropriate.

4.4.5. Conclusions

This study expands on the findings from an earlier pilot study (160) on methyl-donor and cofactor intake during the Gambian rainy season, to provide information on the annual variation in the nutrient intake. Further, dietary intake in this study is coupled with blood biomarkers on each month. The combination of both provides evidence that the rainy season, traditionally considered the nutritionally challenged season, is less deficient in terms of methyl-donor supply. However, the dietary intakes do not always correlate to the plasma status and, from this study it can also be concluded that the biomarkers are more appropriate than the dietary intake to determine the methyl-supply status.

CHAPTER 5: MAIN GROUP

This Chapter describes and discusses the 'main group' study. The main group investigates the association between season of conception and DNA methylation outcomes in offspring. Maternal blood biomarkers of one-carbon metabolism (measured in blood samples collected as close as possible to conception) are hypothesised to explain the effect of season in DNA methylation previously observed (79). The main findings have been written as a paper (Section 5.1.), to be submitted to PLoS Medicine. The remainder of Chapter 5 provides study detail beyond that presented in the prepared manuscript

5.1.RESEARCH PAPER III: 'Maternal one-carbon metabolism and infant DNA methylation: an experiment of nature in rural Gambia'

Article cover sheet:

1. For a 'research paper' prepared for publication but not yet published

   1.1. Where is the work intended to be published?

   This paper presents the main findings of the main group study and is planned to be submitted to PLOS Medicine.

   List the paper's authors in the intended authorship order

   Dominguez-Salas P, Moore SE, da Costa KA, Cox SE, Dyer RA, Fulford AJC, Innis SM, Laritsky E, Zeisel SH, Waterland RA, Prentice AM and Hennig BJ.

   1.2. Stage of publication – **Not yet submitted**/Submitted/Undergoing revision from peer reviewers' comments/In press

   Undergoing revision from co-authors.

2. 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 conceived and designed the study, conducted the research and analysed the data, in collaboration with some of the authors (Moore, Cox, Fulford, Prentice, Waterland, Hennig). I wrote the manuscript and the key co-authors (Moore, Fulford, Prentice, Waterland, Hennig) have provided comments, which were incorporated prior to this thesis submission. The paper is waiting for review by all co-authors before submission to the peer reviewed journal.

Candidate's signature _____

Supervisor or senior author's signature to confirm role as stated in (2) _____

Maternal One-Carbon Metabolism and Infant DNA Methylation: An Experiment of Nature in Rural Gambia

Paula Dominguez-Salas[1,2], Sophie E. Moore[1,2], Kerry-Ann da Costa[3], Sharon E. Cox[1], Roger A. Dyer[4], Anthony J. C. Fulford[1,2], Sheila M. Innis[4], Eleonora Laritsky[5], Steven H. Zeisel[3], Robert A. Waterland[5], Andrew M. Prentice[1,2], Branwen J. Hennig[1]

Affiliations

[1]MRC International Nutrition Group, London School of Hygiene and Tropical Medicine, London, UK

[2]MRC Keneba, MRC Unit, The Gambia

[3]Department of Nutrition, Gillings School of Global Public Health and School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[4]Department of Pediatrics, University of British Columbia, Vancouver, British Columbia, Canada

[5]Departments of Pediatrics and Molecular & Human Genetics, Baylor College of Medicine, USDA/ARS Children's Nutrition Research Center, Houston, Texas, USA

ABSTRACT

Periconceptional maternal supplementation with folic acid, vitamin B12, choline and betaine can induce DNA methylation changes in mice, resulting in phenotypic differences in their offspring. We investigated whether seasonally-driven natural fluctuations in the nutritional status of rural Gambian women, when occurring during the initial stages of embryonic development, can lead to detectable differences in DNA methylation patterns of their infants.

Women (n=167) who conceived at the peak of either the rainy (July-September) or dry (February-April) season were enrolled. Blood samples collected at 8.6 weeks after conception (± 0.12 weeks) were used to assess the concentrations of choline, betaine, folate, methionine and vitamins B2, B6 and B12, as well as S-adenosylhomocysteine (SAM), S-adenosylmethionine (SAH), homocysteine and dimethylglycine (DMG). DNA was obtained from infants born to these women for evaluation of locus-specific DNA methylation at metastable epialleles, by pyrosequencing. Seasonal variation was observed for maternal biomarker levels and infant DNA methylation. The rainy season was associated with higher infant DNA methylation levels (4.92% for *PAX8* and 3.49% for *RBM46*), as well as maternal SAM:SAH ratio (55.8%) and concentrations of most methyl-donors (folate (42.2%), betaine (23.3%) and methionine (12.5%)) and cofactors (B2 and B6). Other maternal biomarkers did not show seasonal variation (choline) or were higher in the dry season (e.g. active B12). DNA methylation at the loci studied was associated with the SAM:SAH and DMG:betaine ratios, with B2 and with cysteine, but not with the rest of the biomarkers.

These data indicate that one-carbon metabolite status varies seasonally (as expected). Observation of differences in DNA methylation pattern by season was replicated and there are associations between maternal biomarker status and infant DNA methylation. Confirmation of a causal relationship requires supplementation trials.

5.1.1. INTRODUCTION

An adverse prenatal environment can programme disease susceptibility in adulthood and such processes are likely regulated – at least in part – by epigenetic mechanisms (19). Epigenetic marks regulate gene expression without altering DNA sequence. Growing evidence is accumulating on the sensitivity of epigenetic marks to external influences such as nutrition (161), and on the importance of epigenetic dysregulation in diseases such as cancer (162) and diabetes (163). It is thought that epigenetics may account for at least some of the missing heritability of complex diseases and act as the interface between early environmental cues and genes. DNA methylation is the best-characterised mechanism of epigenetic modification (110). Methyl-groups ($-CH_3$) originate from diet via one-carbon metabolism (Figure 4.1) and bind covalently to cytosines within cytosine-phosphate-guanine (CpG) dinucleotides. Two complementary pathways exist (folate- and betaine-dependent) to make available methyl groups by remethylating homocysteine into methionine. Yet, how DNA methylation operates and is regulated throughout early development is poorly understood (111). Environmentally induced DNA methylation changes are not limited to the periconceptional period (75), but it has been shown that early after fertilisation, at least in mice, a critical window exists for the establishment of new DNA methylation marks (111). Maternal supplementation of dams with methyl-donors and cofactors has been demonstrated to modify offspring gene expression and hence phenotype (68, 112).

In humans, direct evidence exists for nutritional deficiencies of methyl-donors and cofactors in pregnancy and offspring outcomes such as neural tube defects (164) and orofacial clefts (165), which are thought to be epigenetically mediated. Further, epidemiological evidence from survivors of the Dutch Hunger study suggests that severe famine during the prenatal period can lead to alterations in DNA methylation in the offspring, dependent on offspring sex and gestational age at the time of exposure (74, 75). Moreover, periconceptional folic acid supplementation has been shown to increase DNA

methylation in the IGF2 gene in newborns (166). In The Gambia, micronutrient supplementation during pregnancy was associated with variations in global DNA methylation (104) and DNA methylation at imprinted loci in the offspring (105).

DNA methylation can be cell-specific, thus methylation patterns of one tissue may not reflect systemic DNA methylation (19). So-called metastable epialleles (MEs) are genomic loci at which the epigenotype is established stochastically in the early embryo, prior to gastrulation, and subsequently maintained across all germ-layer cell lineages. Thus, MEs exhibit inter-tissue concordance and allow the use of DNA from easily accessible venous blood for analysis. MEs are thought to be particularly sensitive to environmental exposures, and are therefore suitable candidates for foetal programming studies (167). In rural Gambia, we previously reported an effect of season of conception on DNA methylation at MEs (79). Mean DNA methylation was consistently higher (up to 10% difference) for five genes (*BOLA3*, *FLJ20433*, *PAX8*, *SLITRK1* and *ZFYVE28*) in children conceived during the rainy season months of August to September. We have also reported that seasonal variation exists in dietary intake and metabolic blood biomarkers of methyl-donors in rural Gambia (research paper II, Section 4.1). It remains unclear, however, whether the observed season-of-conception effect on DNA methylation is caused by variation in maternal nutrition and/or by other factors.

The natural pattern of seasonality in rural Gambia provides an experimental setting to explore the role of maternal nutritional status on epigenetic programming in their infants. Here, we aim to: i) replicate and expand on previously identified associations between season of conception and infant DNA methylation; ii) explore whether seasonal differences exist in a comprehensive set of maternal biomarkers of methyl-donor status around the time of conception; and iii) investigate whether such variations could explain the associations between conception season and offspring DNA methylation.

5.1.2. METHODS

5.1.2.1. Study population

This was an observational prospective cohort study, conducted between July 2009 and July 2011 in 34 villages across the rural West Kiang District of The Gambia.

Mothers: All women of reproductive age (18-45 years) registered in the West Kiang Demographic Surveillance System (DSS) were invited to participate. Exclusion criteria included confirmed pregnancy at time of recruitment, menopause, or likely migration (short- or long- term) away from West Kiang. Each month from enrolment, consenting women were visited at the village health post for height and weight measurement (Tanita DH305 scales (Tanita Corporation, Japan) and Leicester height measure (Seca 214, UK)) and to answer a short questionnaire on the date of their last menstrual period (LMP). On the first report of a missed menses, a 10ml fasting venous blood sample was collected for the purpose of blood biomarker assessment. Following a second consecutive missed period the next month, a urine sample was collected for pregnancy testing (this system was set-up to avoid early disclosure of pregnancy). If the test was negative the woman continued to be visited monthly, and her blood sample from the previous month was discarded. If the test was positive, the woman was invited to the MRC Keneba field station for confirmation of pregnancy by ultrasound examination and, where indicated, a full antenatal check. Women who conceived during the peak of the rainy (July-September 2009) or dry (February-April 2010) season and with a maternal blood sample collected within the first 16 weeks from conception were then fully enrolled. Multiple pregnancies were excluded. Recruitment into the study ran in tandem to a supplementation trial on Early Nutrition and Immune Development (the ENID Trial) (ISRCTN49285450). Supplementation (including iron-folate, as per national policy) only started after confirmation of pregnancy, i.e. after sample collection for this study.

Infants*:* Between 3-9 months after delivery, an infant sample of venous blood (3 ml) was collected for DNA extraction by a trained nurse. Based on previous birth rates, we anticipated recruiting up to 100 women-infant pairs in each of the two 3-month peak season time windows.

The Scientific Coordinating Committee of the MRC Unit, The Gambia, granted scientific approval and the joint Gambian Government/MRC Unit The Gambia Ethics Committee (SCC/EC 1151) and the LSHTM Ethics Committee (EC 5525) granted ethical permission. Following community approval in each village, written informed consent was obtained from all study participants and their parents or guardians.

5.1.2.2. Maternal blood biomarkers

Conception date was calculated by adding 14 days to the estimated date of onset of the last menses, based on the gestational age determined by ultrasound at the time of the first antenatal check. The first day of the last menses is estimated to be 14 days prior to fertilisation. A cut-off for 'early pregnancy' defined at 16 weeks from conception at the time of the blood sample collection was implemented, to allow for late detection of pregnancy.

Blood biomarker measurements included folate, B2, B6, B12, active B12 (holotranscobalamin, the biologically active form of B12), choline, betaine and methionine, as well as homocysteine (tHcy), *S*-adenosylmethionine (SAM), *S*-adenosylhomocysteine (SAH) and dimethylglycine (DMG). Maternal samples were collected in the field and transported on ice to the MRC Keneba laboratory for immediate processing. SAM may convert into SAH over time and therefore special care was given to process and freeze the plasma sample without delay. All plasma biomarkers were assessed at the Department of Pediatrics, University of British Columbia, Canada, as previously described (research paper II, Section 4.1). Briefly, SAM, SAH, free choline (it has been observed that high concentrations of free choline are present in foetal plasma (168, 169)), betaine, DMG, total

homocysteine (tHcy), methionine, cysteine and B6 (measured as pyridoxal (PL), pyridoxal-5'-phospahate (PLP) and 4-pyridoxic acid (PA)) were analysed by liquid chromatography-tandem mass spectrometry (LC-MS-MS) as previously reported (114-116). B12 and active B12 and folate were analysed by a microparticle enzyme intrinsic factor assay and by ion capture assay respectively, on an AxSyM analyser (Abbot Laboratories, Chicago, IL). B2 status was determined in red blood cells (RBC) at MRC Human Nutrition Research (HNR), Cambridge, UK, using the erythrocyte glutathione reductase activation coefficient (EGRAC) assay, performed on a microplate. Higher EGRAC values denote B2 deficiency.

5.1.2.3. Infant DNA methylation

DNA was extracted from venous blood using a standard salting-out method (170). Extracted DNA was cleaned using the Chelex-100 (BIO-RAD) protocol (171). DNA methylation analysis was carried out at the Baylor College of Medicine, USDA/ARS Children's Nutrition Research Center, Houston, USA.

Four previously described MEs, namely *BOLA3*, *PAX8*, *EXD3* (*FLJ20433* in the current human genome assembly HG19) and *ZFYVE28* and three newly identified MEs, i.e. *RBM46*, *PARD6G* and *ZNF678*, were investigated. Briefly, these MEs were determined through a custom methyl-specific amplification microarray (MSAM) screening procedure with validation by bisulfite pyrosequencing (172). Loci with tissue-specific methylation, and CpG affected by polymorphisms that could confound the MSAM detection scheme, were not considered (79). CpG-specific methylation in the current infant DNA samples was measured by quantitative bisulfite pyrosequencing (Pyro Gold reagents and a PSQTM HS 96 pyrosequencer, from Biotage), as described elsewhere (172). Briefly, 0.5-2 μg of DNA was bisulfite treated, followed by locus-specific PCR amplification and pyrosequencing to measure methylation at 4 to 12 CpG sites per locus (see **Table 5.1** for details). Each pyrosequencing assay covered 50-70bp of DNA and was validated by inclusion of 0, 25, 50, 75 and 100% methylated standard human DNA samples.

Table 5.1: Details for the PCR and Pyrosequencing conditions for DNA methylation analyses [79]

Locus name and location (chromosomal position based on human genome build 19)	**Number of CpG sites**	**Forward Primer**	**Reverse Primer**	**Product size (bp)**	**Annealing temperature (°C)**	**Pyrosequencing Primer**
BOLA3 (5kb downstream) chr2:74,357,770	5	GGTGTATTTAAGTAT AGAGAAGGTGGAGAT	Btn-AAAACAACTAAAACTCA CAAACCACTACTA	206	61	GTGGAGATGGAGGGA
PAX8 (Intron 9) chr2:113,992,770	5	3': GGGGTGGATGAGATT GAGGTTAGA	3': Btn-CCTTCAATACCTTTCCCC ATACTACC	171	67	3': GGTTTGTTTTGAGGAT
RBM46 (Intron 1) chr4:155,703,000	12	3': TTGTATGGTGAGGGT TTAG	3′: TCTAAAACCAAACTACT AAATCT	293	51	3′: GTGTTATTTTTTTGATA
ZFYVE28 (Promoter CGI) chr4:2,366,687	5	5′: TTTAGTAGGGGYGGYG TAGTTTTAGTTATA	5′: Btn-AAACCTAACRCCTAAAA AATAACC	82	59	5′: GGYGTAGTTTTAGTTATAGAGT
EXD3 (Intron 1) chr9:140,312,195	5	Btn-AGGTTATTGAGTTGG GTTTTTT	CTCCTCAAATCCTCAAA CTCTATCC	134	53	CCTCAAACTCTATCCTTTC
PARD6G (Last exon) chr18:77,918,167	6	5′: GTAGATGGAATAGTA GTGTTAGGTGTATGA	5′: TCATTCACAACCAACAA CC	437	53	5′: GGGTTGGAGAAGGTG
ZNF678 (5kb upstream) chr1:227,746,190	4	5′: GGGTGTTGAAGGTTT TTT	5′: CAAACCAAACTCCTAAT ACTAT	220	55	5′: CTACCTAATAATACTACAATC

bp: base-pairs; chr: chromosome; CpG: cysteine-phosphate-guanine

5.1.2.4. Statistic analysis

All analyses were performed in Stata 11.0 (StataCorp, College Station, TX). The primary exposure was season of conception (rainy *vs.* dry season) and DNA methylation the outcome measure. Maternal blood concentrations of one-carbon metabolites acted as both an outcome of season and a intermediate exposure for DNA methylation.

Study population characteristics: Demographic characteristics within both seasonal groups were compared using the Wilcoxon-Mann-Whitney non-parametric test (for continuous variables) and Chi-square test (for categorical variables). Body mass index (BMI) was calculated in kg/m^2. The infant Z-scores of weight-for-length (WLZ) and length-for-age (LAZ) were calculated using the 2006 WHO Growth Standards (173).

Maternal blood biomarkers: In addition to the 15 individual biomarkers under investigation, two further variables were created: the SAM:SAH and the DMG:betaine ratios, as measures of methylation activity and betaine-pathway activity, respectively. All blood biomarkers showed signs of a positively skewed distribution and were therefore logarithmically transformed and reported as geometric means. Because it would have been impossible to draw blood samples timed to coincide with the immediate post-conception days we enrolled a separate group of non-pregnant women (the indicator group) and studied their diets and plasma biomarker status at monthly intervals (full details described in research paper II, Section 4.1). This allowed us to extrapolate the biomarker levels of the pregnant women back to the levels likely to have been present at conception. To achieve the adjustment we first assessed the accuracy of the autocorrelation in the indicator group (reliability and stability of each biomarker) and then tested how much the seasonal patterns (Fourier series terms, namely sin1 cos1 sin2 cos2) matched between the indicator group and pregnant women by fitting the same seasonal patterns (research paper II, Section 4.1). Then we used the non-pregnant women seasonal data for the back extrapolation, according to the following equation:

$$\log\_y_{adj} = \log\_y + \text{mean (at conception)}_{SInd} - \text{mean (at measurement)}_{SInd}$$

y_{adj}: outcome variable adjusted for weeks since conception and seasonality
$Mean_{SInd}$: mean of Seasonal non-pregnant women (Indicator) group

Differences in maternal biomarker concentration were tested by comparing both seasonal groups of pregnant women (rainy *vs*. dry season conception) with one-way ANOVA. Potential predictors of blood biomarker concentration (e.g. age, BMI) were determined by multiple regression analysis. This was conducted on the biomarker data prior to back-extrapolation, to explore the independent effects of season and weeks since conception. Correlations were calculated amongst (unadjusted) biomarkers to assess their interdependence.

Infant DNA methylation: Loci that had more than one CpG site failing the quality control cut-off, and individuals with more than one ME that failed to be screened, were excluded from the analysis. Methylation at the majority of CpG sites showed signs of a skewed distribution and hence a logit transformation was performed. For the seasonal comparison of methylation at each ME, multivariate analysis of variance (MANOVA) was used, including all existing CpG sites for each locus. The mean methylation at each locus was then calculated. Correlations were also assessed between mean DNA methylation at each ME to assess their interdependence. Finally, a multilevel Generalised Estimated Equations (GEE) model analysis was used to explore the association between the different non-nutritional predictors (e.g. infant sex or ethnicity) and mean methylation at all MEs jointly, standardised (by dividing by the standard deviation for its variation between women). This multilevel model assumes independence between the variables. Therefore, given that correlation between the different MEs exists, an unstructured correlation matrix was used to fit the model (i.e. with unconstrained conditions in the correlation).

Interrelationships between maternal blood biomarkers and infant DNA methylation: An equivalent multilevel analysis model was used to explore the association between each individual maternal biomarker and methylation at all standardised MEs jointly. All multilevel GEE regressions were adjusted by infant sex and maternal BMI. This model was also used to examine whether the biomarker effects differed by ME. Given the interrelations that can exist between the different metabolites, principal component analysis (PCA) was used to allow compressing of biomarker data into 'component' variables. This reduces the number of dimensions and simplifies the interrelationships between biomarkers, without important loss of information, for the analyses of the effect all biomarkers on DNA methylation. To account for the multiple testing, a Bonferroni correction was applied, with a cut-off of significance of 0.0005 (p-value = 0.05/(13 biomarkers x 7 loci)).

5.1.3. RESULTS

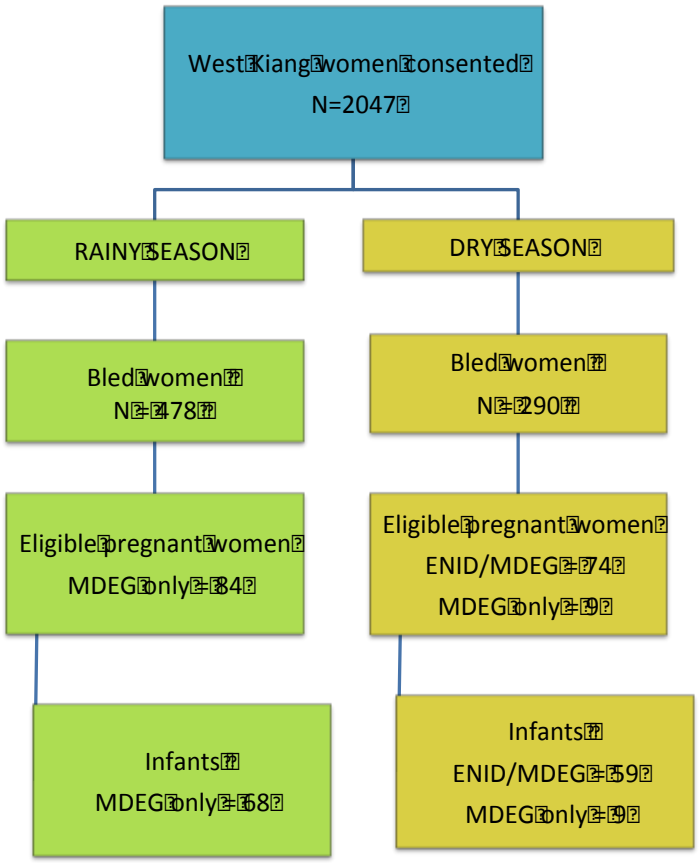
5.1.3.1. Study population

Mothers: 2047 women were consented and visited monthly. During the follow-up period, 335 participants were excluded due to self-reported onset of menopause not identified at consent, out-migration or self-withdrawal. The total number of women conceiving during the *a priori* selected months and having a blood sample collected during the first 16 weeks of pregnancy from conception was 167, recruited across 24 villages. Of these, 84 women (50.0%) conceived during the rainy season. The mean age of all the women was 29.2 years, the majority of which were of self-reported Mandinka ethnicity (86.2%), the remainder being Fula or 'Other'. The average BMI of women, measured at a mean gestational age of 8.6 weeks from conception (at the time of the blood collection) was 21.2kg/m^2 (range 15.08- 33.45).

Infants: 167 infants were recruited prenatally, but only 136 of these made it to follow up; the rest were lost through miscarriage, infant death, difficulty to collect blood sample or withdrawal (see **Figure 5.1**). The majority of infants were of maternally-reported Mandinka ethnicity (80.7%) and 52.9% were male. The mean (SD) age of infants at bleeding was 3.5 (+/-0.98) months (range 2.0-8.2 months). The mean (SD) WLZ was -0.43 (+/-1.32) and the LAZ was -0.55 (+/-1.01).

The summary statistics of the study sample groups are provided in **Table 5.2**. No significant differences in demographic characteristics were found for either mothers or infants by season.

Figure 5.1: Participants enrolment



MDEG: Methyl-Donors and EpiGenetics study (current study (SCC/EC 1151))
ENID: Early Nutrition and Immune Development trial (ISRCTN49285450)
Rainy season (in green): July-September 2009; Dry season (in yellow): February-April 2010; Inclusion and exclusion criteria are described in Section 5.1.2.1
Maternal withdrawals N=335
Reasons for loss to follow up of infants: miscarriage (n=6), inability to get infant blood (n=9) or other reasons including infant death, participants moving-away, self-withdrawal (n=7). Some infants (n=9) were bled but the extracted DNA quality was not adequate for the DNA methylation analysis.

Table 5.2: Baseline characteristics of study population by season

	Rainy season		Dry season		
Characteristics	**N** (%)	**Mean** (95% range**)**	**N** (%)	**Mean** (95% range)	**P-value of the difference between seasons**
Mothers					
Age, years	84 (50.3%)	29.4 (15.4-43.3)	83 (49.7%)	29.1 (16.5-41.7)	0.960[a]
Ethnicity* Mandinka Fula Other	 52 (81.3%) 11 (17.2%) 1 (1.6%)	-	 48 (92.3%) 4 (7.7%) 0 (0%)	-	0.200[b]
Weeks from conception	84 (50.3%)	8.4 (-0.3-17.1)	83 (50.3%)	8.9 (7.4-16.3)	0.546[a]
BMI at bleeding	73 (47.4%)	21.2 (14.8-3.2)	81 (52.6%)	21.1 (14.6-27.6)	0.546[a]
Infants					
Males	39 (57.4%)		33 (48.5%)		0.303[b]
Age, months	65 (53.3)	3.9 (1.7-6.1)	57 (46.7%)	3.1 (1.9-4.4)	0.546[a]
Ethnicity: Mandinka Fula Other	 47 (73.4%) 11 (17.1%) 6 (14.1%)		 44 (88%) 3 (6%) 3 (6%)		0.254[b]
Weight for length (Z-score)	64 (53.3%)	-0.4 (-3.4-2.5)	56 (46.7%)	-0.4 (-2.6-1.8)	0.979[a]
Length for age (Z-score)	64 (54.2)	-0.7 (-2.7-1.2)	54 (45.8%)	-0.3 (-2.4-1.7)	0.045[a]

*Ethnicity only available for women with infants; 95% range = Mean ± 2SD
[a]Wilcoxon-Mann-Whitney non-parametric test; [b]Chi-square test

5.1.3.2. Maternal blood biomarkers

Plasma samples were frozen on average 179 minutes after collection in the field (range 38 to 380 minutes). The effect of time of processing (between sample collection and storage in the freezer) was assessed and found to affect DMG, which decreased (β= -0.191 µmol/l per hour, p-value=0.034) and SAH, which increased (β= 0.880 nmol/l per hour, p-value=0.001) with processing time. Linear regressions were used to correct these two variables to time zero. Seasonal patterns of biomarker levels in the pregnant women were tested against our non-pregnant women group to assess the degree of appropriateness of the back-extrapolation of biomarker level to time of conception. The age of this 'indicator' group women ranged from 18 to 45 years (mean (SD) 31.2 (8.0) years). Folate, choline, betaine, B12, Active B12, SAM and the DMG/betaine ratio showed similar patterns in pregnant women compared to the non-pregnant

group (p-value >0.05), whilst methionine, DMG, SAH, B2, B6 vitamers (PL, PLP and PA), SAM:SAH ratio, homocysteine and cysteine differed significantly (p-value <0.05). Seasonal patterns for both the non-pregnant and the pregnant women are shown in **Figure 5.2**.

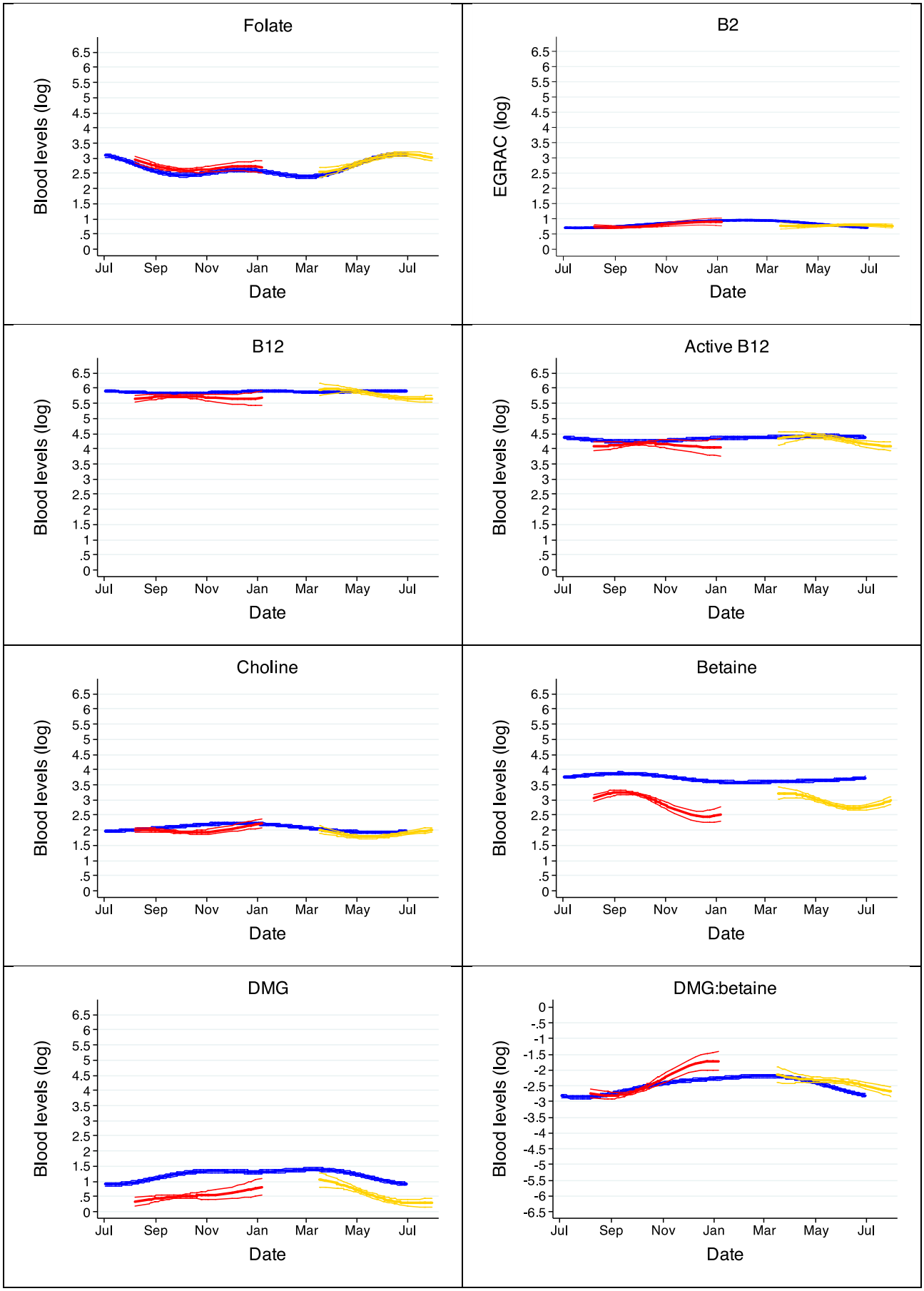
The differences in blood biomarker concentrations between women who conceived during the rainy and the dry season prior to (unadjusted data) and after the back-extrapolation for seasonal trends (based on the data from the group of non-pregnant women) are shown in **Tables 5.3** and **5.4**. After back-extrapolation, geometric mean levels of the following biomarkers were higher in the rainy season when compared with the dry season: SAM:SAH ratio (55.8%), betaine (23.3%), folate (42.2%), methionine (12.5%), PL (41.9%) and PA (29.1%). EGRAC was lower (19.1%), denoting a higher B2 status. Conversely, plasma concentrations of DMG (39.1%), SAH (34.4%), active B12 (17.8%), homocysteine (17.7%) and the DMG:betaine ratio (53.3%) were lower during the rainy than during the dry season. No significant differences were observed between seasons for the remainder of the biomarkers.

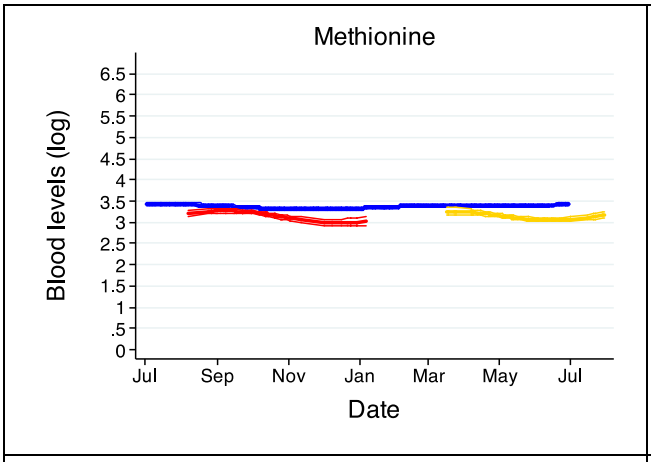
As illustrated in Figure 4.1, one-carbon metabolites are highly interdependent. Pairwise univariate correlations between biomarkers (unadjusted data) were examined and the correlation matrix is shown in **Table 5.5**. It was observed that homocysteine correlated negatively with markers of the folate pathway (r=-0.374 for folate, r=-0.117 for B12 and r=0.195 for B2 status (positive correlation with EGRAC)), but positively with betaine (r=0.308), choline (r=0.177) and DMG (r=0.330) whilst SAM:SAH ratio correlated negatively with both pathways (r=-0.123 for folate, r=-0.224 for choline and r=0.330 for DMG. Looking at correlations between the folate dependent and independent pathways, concentrations of DMG were negatively correlated with folate (r=-0.178) and with B2 (positively correlated with EGRAC, r= 0.120).

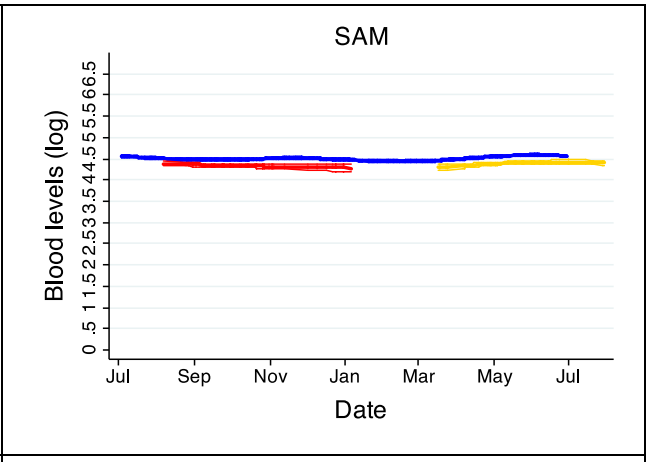
Predictors of biomarker levels were assessed by multiple regression (of unadjusted logged outcomes), including maternal age, BMI, weeks since conception at bleeding and season (sin1, cos1, sin2, cos2) as continuous variables, as well as infant sex (binary variable). Maternal age appeared to affect homocysteine (coef. β=0.016 μmol/l per year, p-value<0.0001) and cysteine (coef β=-0.011 μmol/l per year, p-value=0.027), but not any of the other biomarkers.

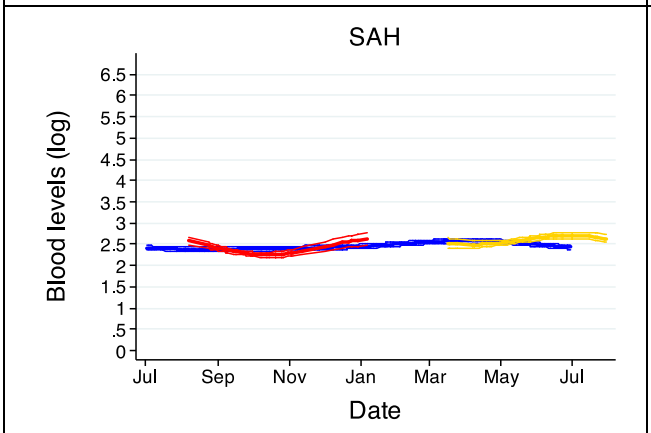
BMI was significantly associated with blood concentrations of B12 (coef β=-0.031 pmol/l per year, p-value=0.008) and active B12 (coef β= -0.041 pmol/l per year, p-value=0.007) only. Weeks since conception (i.e. pregnancy stage) at bleeding had an effect on several of the blood biomarkers, as shown in Table 5.3. There was no association between infant sex and maternal biomarkers, except for SAM (coef β= 0.075 nmol/l per year, p-value=0.018). 71.9% of women reported to have maintained an overnight fast prior to bleeding. However, fasting status did not affect the biomarker concentrations (data not shown).

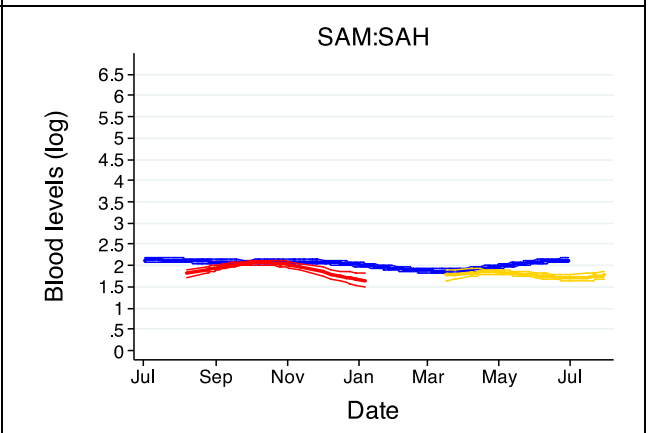
Figure 5.2: Comparison of seasonal patterns in the blood biomarker concentrations between the 'indicator group' and the pregnant women.

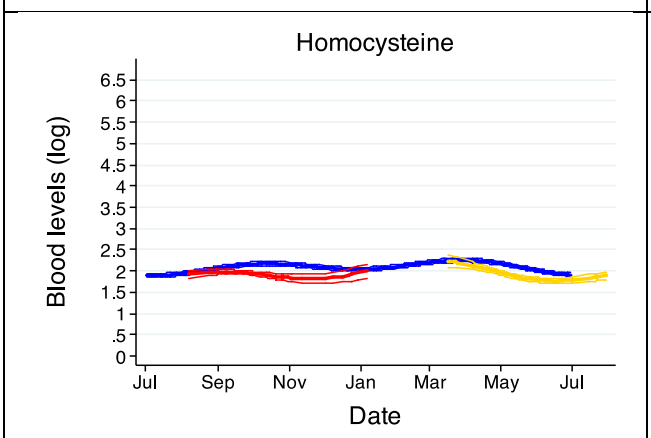


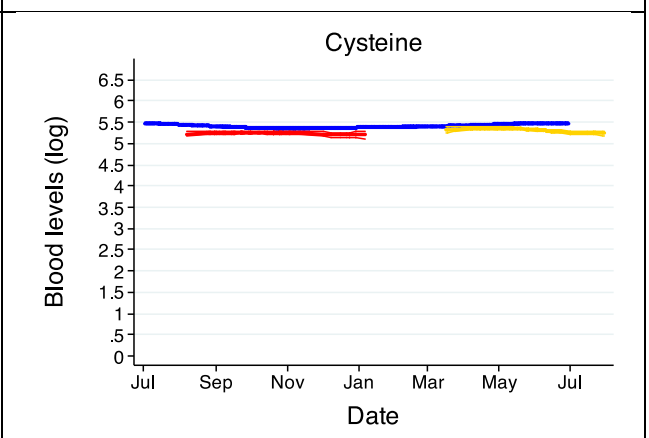


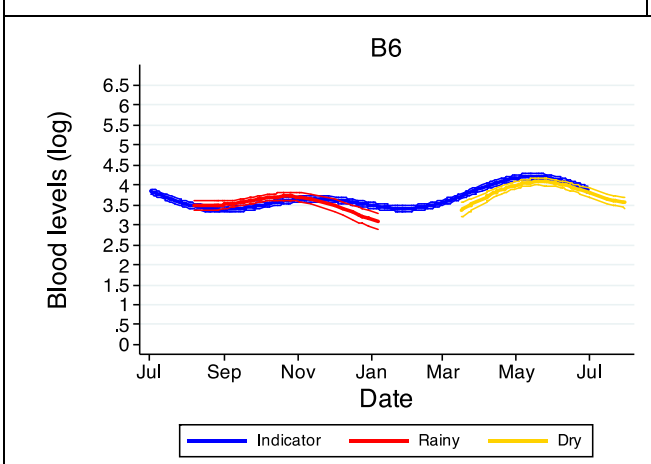












Thick line: geometric mean of the biomarker
Thin lines: 95%CI
Indicator: annual trends for the group of non-pregnant women (n=30)
Rainy/dry: groups of pregnant women enrolled in the study (n=167)
Variables were logarithmically transformed (unadjusted data)
Study period: July 2009- July 2010
Statistical test of the difference between groups:
p-value <0.05 indicates significant difference: in B2, DMG, methionine, SAH, SAM:SAH ratio, homocysteine, cysteine, B6.
p-value>0.05 indicates non-significant difference: folate, B12, active B12, choline, betaine, DMG/betaine ratio and SAM.

Table 5.3: Differences in blood biomarker concentration between women conceiving during the rainy *vs.* dry season, with adjustment for time of bleeding (approximation to conception time), and effect of the weeks of gestation on the biomarker levels

Biomarker	N	Rainy season GM[a] adjusted (95%CI)	Dry season GM[a] adjusted (95% CI)	Between season difference adjusted	Effect of time (weeks) since conception on biomarker levels unadjusted β coef (p-value)[b]
Folate [nmol/l]	163	18.47 (16.81-20.29)	12.99 (12.04-14.01)	5.48*	0.023 (0.020)*
B12 [pmol/l]	164	307.73 (280.73-337.33)	335.04 (306.56-366.17)	-27.31	-0.034 (0.004)*
Active B12 [pmol/l]	164	63.83 (57.19-71.24)	77.64 (70.26-85.80)	-13.81*	0.004 (0.760)
B2 deficiency$_1$ [EGRAC coef]	155	1.98 (1.87-2.09)	2.44 (2.32-2.57)	-0.47*	-0.002 (0.323)
Choline [μmol/l]	163	6.62 (6.23-7.03)	6.86 (6.39-7.36)	-0.24	-0.002 (0.807)
Betaine [μmol/l]	164	21.21 (19.13-23.52)	17.21 (15.21-19.46)	4.00*	-0.094 (<0.0001)*
DMG [μmol/l]	164	1.83 (1.63-2.05)	3.00 (2.73-3.30)	-1.17*	-0.037 (0.002)*
DMG:betaine	164	0.08 (0.07-0.09)	0.18 (0.16-0.20)	-0.10*	0.057 (<0.0001)*
Methionine [μmol/l]	164	25.47 (24.33-26.66)	22.65 (21.74-23.59)	2.82*	-0.010 (0.068)
SAM [nmol/l]	163	77.48 (74.59-80.47)	75.77 (72.91-78.76)	1.70	-0.013 (0.010)*
SAH [nmol/l]	163	7.55 (6.97-8.19)	11.50 (10.50-12.60)	-3.95*	-0.021 (0.053)
SAM:SAH	163	10.26 (9.59-10.98)	6.59 (6.06-7.16)	3.67*	0.009 (0.367)
Homocysteine [μmol/l]	163	6.18 (5.76-6.62)	7.50 (7.03-8.01)	-1.33*	-0.031 (<0.0001)*
Cysteine [μmol/l]	164	199.11 (193.32-205.07)	197.18 (191.27-203.26)	1.93	-0.019 (<0.0001)*
PL [nmol/l]	164	4.74 (4.46-5.04)	3.34 (2.99-3.73)	1.40*	-0.007 (0.522)
PLP [nmol/l]	164	22.69 (20.77-24.80)	25.63 (22.64-29.02)	-2.94	-0.022 (0.052)
PA [nmol/l]	164	9.30 (8.75-9.88)	7.21 (6.35-8.17)	2.09*	0.006 (0.590)
SAM:Hcy	163	12.61 (11.65-13.66)	10.05 (9.34-10.82)	2.56*	0.019 (0.045)*

[a]GM: geometric mean back-extrapolated with time of bleeding; [b]Raw data (non-back-extrapolated), log-transformed adjusted by BMI, time to processing, seasonality (sin1, cos1, sin2, cos2) and fasting; * p-value for significance <0.05

[1]B2 def: B2 deficiency as assessed by EGRAC; DMG, dimethylglycine; Hcy, homocysteine; PA, 4-pyridoxic acid; PL, pyridoxal; PLP, pyridoxal-5'-phosphate; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine.

Table 5.4: Differences in blood biomarker concentration between women conceiving during the rainy vs dry season, without back-extrapolation for time of bleeding

Biomarker	N	Rainy season GM[a] unadjusted (95%CI)	Dry season GM[a] unadjusted (95%CI)	Between season difference unadjusted data
Folate [nmol/l]	163	14.41 (13.33-15.56)	18.69 (17.21-20.30)	-4.28*
B12 [pmol/l]	164	300.13 (273.91-328.87)	337.37 (309.08-368.24)	-37.24
Active B12 [pmol/l]	164	61.93 (55.67-68.89)	78.33 (70.86-86.59)	-16.4*
B2 deficiency[1] [EGRAC coef]	155	2.17 (2.04-2.29)	2.17 (2.06-2.28)	0
Choline [μmol/l]	163	7.37 (6.95-7.82)	6.35 (5.92-6.82)	1.02*
Betaine [μmol/l]	164	20.92 (18.75-23.35)	18.31 (16.27-20.62)	2.61
DMG [μmol/l]	164	1.71 (1.53-1.91)	1.82 (1.59-2.07)	-0.11
DMG:betaine	164	0.10 (0.09-0.11)	0.13 (0.12-0.15)	-0.03*
Methionine [μmol/l]	164	24.11 (23.01-25.26)	22.92 (22.02-23.87)	1.19
SAM [nmol/l]	163	75.97 (73.21-78.83)	81.35 (78.38-84.48)	-5.38*
SAH [nmol/l]	163	10.43 (9.85-11.05)	13.64 (12.74-14.59)	-3.21*
SAM:SAH	163	10.09 (9.44-10.79)	7.41 (6.85-8.03)	2.67*
Homocysteine [μmol/l]	163	6.94 (16.54-7.36)	6.67 (6.23-7.15)	0.27
Cysteine [μmol/l]	164	189.22 (183.57-195.05)	205.31 (199.43-211.37)	-16.09*
PL [nmol/l]	164	4.45 (4.29-4.62)	4.29 (3.88-4.75)	0.15
PLP [nmol/l]	164	21.71 (20.10-23.46)	32.32 (28.84-36.21)	-10.60*
PA [nmol/l]	164	9.63 (9.21-10.07)	9.31 (8.21-10.55)	0.33
SAM:Hcy	163	10.97 (10.26-10.72)	12.20 (11.24-13.26)	-1.24*

[a]GM: geometric mean non-back-extrapolated with time of bleeding;; * p-value for significance <0.05

[1]B2 def: B2 deficiency as assessed by EGRAC. DMG, dimethylglycine; Hcy, homocysteine; PA, 4-pyridoxic acid; PL, pyridoxal; PLP, pyridoxal-5'-phosphate; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine

Table 5.5: Pairwise univariate correlations between maternal biomarkers

Biomarker	Folate	B12	Active B12	B2 def[1]	Betaine	Choline	DMG	DMG: betaine	Met	SAM	SAH	SAM: SAH	Hcy	Cys	PL	PLP
B12	-0.141**	1.000														
Active B12	-0.057	0.503**	1.000													
B2 EGRAC	-0.053	-0.181**	-0.148**	1.000												
Betaine	-0.118**	0.108**ǫ	-0.160**	-0.059	1.000											
Choline	-0.070*	-0.200**	-0.202**	0.043	0.372**	1.000										
DMG	-0.178**	0.094**	0.039	0.120**	0.397**	0.188**	1.000									
DMG:betaine	-0.002	-0.042	0.226**	0.092**	-0.674**	-0.294**	0.410**	1.000								
Met	-0.176**	0.012	0.023	-0.163**	0.369**	0.216**	-0.035	-0.370**	1.000							
SAM	0.173**	0.030	0.095**	-0.155**	0.119**	0.075*	0.079*	-0.031	0.072*	1.000						
SAH	0.209**	0.053	0.033*	-0.099**	0.007	0.248**	0.101**	0.075	0.063*	0.487**	1.000					
SAM:SAH	-0.123**	0.019	0.008	-0.002	0.046	-0.224**	-0.082*	-0.118**	-0.009	0.169**	-0.751**	1.000				
Hcy	-0.374**	-0.117**	-0.047	0.195**	0.308**	0.177**	0.330**	-0.044	0.191**	-0.026	0.142**	-0.188**	1.000			
Cys	0.131**	0.200**	0.242**	-0.014	0.278**	0.058	0.192**	-0.096**	0.186**	0.056	0.186**	0.148**	0.535**	1.000		
PL	0.132**	-0.016	-0.058*	0.065	0.081*	0.068	0.031	-0.046	0.057	0.225**	-0.080**	0.083*	0.060*	0.217**	1.000	
PLP	0.338**	0.072*	0.120**	-0.014	0.118**	-0.157**	0.055	-0.013	0.065*	0.008	0.093**	-0.013	0.009	0.392**	0.641**	1.000
PA	0.264**	-0.082*	0.044	-0.034	-0.091**	0.005	-0.025	0.075*	0.140**	0.238**	0.138**	-0.003	0.146**	0.071*	0.574**	0.510**

Coefficient correlations and significance are shown *P<0.05; **P<0.005.Variables logarithmically transformed (unadjusted data), n=164[1]Higher values of EGRAC B2 measurement indicate deficiency of B2. Cys, cysteine; DMG, dimethylglycine; Hcy, homocysteine; Met, methionine; PA, 4-pyridoxic acid; PL, pyridoxal; PLP, pyridoxal-5'-phosphate; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine

5.1.3.3. Infant DNA methylation

Methylation at mean gene DNA methylation at each locus was compared by season of conception, and showed consistent trends for increased methylation in infants conceived during the rainy season, however not all these associations reached statistical significance (see **Table 5.6**). Mean methylation was significantly different between season only for *RBM46* and *PAX8*. The same trend was observed in individual CpG sites (**Table 5.7**)**.** The seven loci investigated were observed to have different degrees of mean methylation, ranging from low (29.73% *ZFYVE28*) to high (80.85% for *EXD3*). Similarly, we observed substantial variation in the proportion of CpG site methylation within the same locus, particularly for *BOLA3* (range 27.9%-82.6%), as shown in Table 5.7. Within each locus, methylation of individual CpG sites was significantly and strongly correlated ($r \geq 0.6$), for all loci with the exception of ZNF678 (see **Tables 5.8 A-G**). The correlations between mean methylation at each locus were examined in **Table 5.9**. Only 65% were significant and the effect size observed was modest ($r<35\%$ in all cases). With multilevel regression, it was observed that there was no interaction between mean methylation at different genes and season, and no statistically significant difference in effects between MEs. Other predictors of infant DNA methylation were analysed by multiple regression multilevel analysis. Maternal BMI at bleeding had a negative effect on methylation (β=-0.044 per %methylation/BMI unit, p-value=0.007). Infant sex did not have a significant effect for all MEs analysed together, but when MEs were looked at individually (for all CpGs within a loci) with MANOVA, a higher methylation was observed for *RBM46* (β= -0.193, p=0.033) and *ZNF678* (β=-0.242, p=0.009) in females and higher methylation in males for *EXD3* (β= 0.245, p=0.016). No significant interactions were found between sex and season of conception (data not shown) and there was no association between infant DNA methylation and maternal age or ethnicity.

Table 5.6: Mean DNA methylation at metastable epialleles and season of conception

Locus (N CpG site)	**Rainy season mean methylation** [%] (95% CI)	**Dry season mean methylation** [%] (95% CI)	**Between season difference** Δ%[a] (SD)	**Effect size** SD units[b]	**P-value**
BOLA3 (5)	51.37 (48.12-54.62)	50.46 (46.65-54.28)	0.9 (14.05)	0.064	0.2263
PAX8 (5)	76.79 (75.04-81.92)	71.87 (66.50-77.25)	4.92 (17.68)	0.278	0.0099*
RBM46 (12)	71.56 (69.58-73.54)	68.07 (65.56-70.57)	3.49 (8.92)	0.391	0.0104*
ZFYVE28 (5)	29.16 (26.03-32.29)	25.29 (22.07-28.50)	3.61 (12.64)	0.286	0.2692
EXD3 (5)	80.85 (78.95-82.76)	79.13 (77.12-81.14)	1.72 (7.80)	0.220	0.4008
PARD6G (6)	53.49 (50.24-56.74)[1]	48.66 (44.27-53.05)[1]	4.83 (14.73)	0.328	0.2055
ZNF678 (4)	31.68 (29.55-33.81)	29.73 (27.28-32.18) [3]	1.95 (9.06)	0.215	0.1148

[a] Mean absolute difference in DNA methylation between dry and rainy season; [b] Effect size: observed difference in DNA methylation divided by the SD; * p-value < 0.05
N = 67 for rainy season and n=61 for dry season unless otherwise indicated
[1]N rainy season= 65; N dry season =56; [2]N rainy season= 66; [3]N dry season= 60

Table 5.7: DNA methylation at individual CpG sites of the metastable epialleles analysed and season of conception

Locus CpG site	**Rainy season mean methylation** [%] (95%CI)	**Dry season mean methylation** [%] (95%CI)	**Between season difference** Δ%[a](SD)	**Effect size** SD units[b]	**P-value[c]**
BOLA3 CpG1	58.28 (54.41-62.14)	57.91 (53.40-62.43)	0.36 (16.65)	0.022	0.9026
BOLA3 CpG2	82.56 (80.53- 84.58)	81.92 (79.51-84.32)	0.64 (8.81)	0.073	0.6832
BOLA3 CpG3	51.13 (47.24-55.011)	50.29 (45.94-54.64)	0.83 (16.38)	0.051	0.7747
BOLA3 CpG4	27.96 (24.91-31.02)	26.24 (22.46-30.02)	1.73 (13.58)	0.127	0.4757
BOLA3 CpG5	36.91 (32.92- 40.90)	35.96 (31.12-40.79)	0.95 (17.54)	0.054	0.7600
PAX8 CpG1	76.79 (73.35-80.23)	69.74 (64.36-75.13)[3]	7.05 (17.70)	0.398	0.0266*
PAX8 CpG2	79.09 (75.65-82.53)	71.98 (66.51-77.44)	7.12 (17.78)	0.401	0.0261*
PAX8 CpG3	79.69 (76.21-83.16)	72.48 (66.80-78.16)	7.2 (18.41)	0.391	0.0292*
PAX8 CpG4	78.13 (74.57-81.68)	71.31 (65.84-76.78)	6.82 (18.08)	0.377	0.0355*
PAX8 CpG5	78.69 (75.13-82.26)	73.99 (68.67-79.31)	4.7 (17.77)	0.264	0.1382
RBM46 CpG1	78.30 (76.20-80.39)	74.25 (71.38-77.12)	4.05 (9.90)	0.409	0.0228*
RBM46 CpG2	79.89 (78.24-81.53)	76.81 (74.68-78.93)	3.08 (7.49)	0.411	0.0222*
RBM46 CpG3	76.37 (74.18-78.56)	74.59 (71.81-77.38)	1.78 (9.89)	0.180	0.3135
RBM46 CpG4	71.28 (68.74-73.81)	66.04 (62.72-69.36)	5.24 (11.65)	0.450	0.0125*
RBM46 CpG5	70.62 (68.43-72.80)	66.79 (63.89-69.69)	3.83 (10.13)	0.378	0.0352*
RBM46 CpG6	81.44 (79.06-83.81)	78.48 (75.60-81.35)[3]	2.96 (10.38)	0.285	0.1119
RBM46 CpG7	73.45 (70.90-76.01)	68.08 (65.16-71.00)	5.38 (10.89)	0.494	0.0063*
RBM46 CpG8	72.14 (69.95-74.32)	66.90 (64.20-69.61)	5.23 (9.73)	0.538	0.0030*
RBM46 CpG9	72.34 (70.06-74.61)	68.55 (65.67-71.43)	3.79 (10.25)	0.370	0.0394
RBM46 CpG10	75.43 (73.01-77.85)	71.91 (68.91-74.90)	3.52 (10.77)	0.327	0.20678
RBM46 CpG11	50.93 (48.89-52.98)	49.12 (46.63-51.62)	1.81 (9.02)	0.201	0.2604
RBM46 CpG12	56.56 (54.74-58.31)	55.47 (53.40-57.55)	1.05 (7.67)	0.137	0.4405

Locus CpG site	**Rainy season mean methylation** [%](95%CI)	**Dry season mean methylation** [%] (95%CI)	**Between season difference** Δ%[a](SD)	**Effect size** SD units[b]	**P-value[c]**
ZFYVE28 CpG1	30.27 (26.65-33.89)	25.55 (21.95-29.15)	4.72 (14.41)	0.328	0.0677
ZFYVE28 CpG2	28.05 (24.80-31.30)	24.09 (20.41-27.78)	3.95 (13.80)	0.286	0.1091
ZFYVE28 CpG3	36.91 (33.06-40.75)	30.58 (26.54-34.61)	6.33 (15.70)	0.403	0.0250*
ZFYVE28 CpG4	26.12 (23.22-29.01)	24.43 (21.67-27.19)	1.69 (11.32)	0.149	0.4030
ZFYVE28 CpG5	24.45 (21.00-27.92)	21.78 (17.82-25.74)	2.67 (14.77)	0.181	0.3107
EXD3 CpG1	85.01 (83.06-86.95)	82.50 (80.35-84.66)	2.51 (8.15)	0.308	0.0863
EXD3 CpG2	77.76 (76.26-79.26)	76.22 (74.63-77.81)	1.54 (6.15)	0.250	0.1610
EXD3 CpG3	82.33 (80.307-84.36)	80.47 (78.25-82.69)	1.86 (8.45)	0.220	0.2165
EXD3 CpG4	78.02 (75.47-80.58)	77.05 (74.82-79.29)	0.97 (9.64)	0.101	0.5722
EXD3 CpG5	81.13 (79.18-83.08)	79.40 (77.32-81.47)	1.73 (8.01)	0.216	0.2250
PARD6G CpG1	60.25 (56.94-63.56)[1]	57.01 (52.58-61.44)[1]	3.24 (14.90)	0.217	0.2354
PARD6G CpG2	50.98 (47.70-54.26)[1]	46.05 (41.43-50.67)[1]	4.93 (15.23)	0.324	0.0782
PARD6G CpG3	55.91 (52.47-59.34)[1]	49.23 (44.21-54.25)[1]	6.68 (16.33)	0.409	0.0265*
PARD6G CpG4	51.11 (47.47-54.74)[1]	46.75 (42.10-51.37)[1]	4.37 (15.92)	0.274	0.1352
PARD6G CpG5	57.66 (54.37-60.94)[1]	52.72 (48.29-57.15)[1]	4.94 (14.87)	0.332	0.0710
PARD6G CpG6	45.03 (41.14-48.92)[1]	39.34 (34.65-44.04)[1]	5.69 (16.45)	0.346	0.0623
ZNF678 CpG1	36.08 (33.49-38.66)	36.12 (33.67-38.58)[3]	-0.04 (10.05)	-0.004	0.9798
ZNF678 CpG2	24.58 (22.30-26.85)	22.83 (19.96-25.69)[3]	1.75 (10.17)	0.172	0.3365
ZNF678 CpG3	26.75 (24.37-29.12)	23.66 (21.08-26.24)[3]	3.09 (9.82)	0.315	0.0808
ZNF678 CpG4	39.51 (36.91-42.11)[2]	36.31 (33.39-39.22)[3]	3.20 (10.87)	0.294	0.1027

[a] Mean absolute difference in DNA methylation between dry and rainy season; [b] Effect size: observed difference in DNA methylation divided by the SD in the rainy season; [c] One-way ANOVA; * p-value < 0.05

N = 67 for rainy season and n=61 for dry season unless otherwise indicated; [1]N rainy season= 65; N dry season =56; [2]N rainy season= 66; [3]N dry season= 60

Table 5.8 A to G: Correlation coefficients between CpG sites within each metastable epiallele methylation percentage

A) *BOLA3*

BOLA3	**CpG1**	**CpG2**	**CpG3**	**CpG4**	**CpG5**
CpG1	1.000				
CpG2	0.921	1.000			
CpG3	0.932	0.860	1.000		
CpG4	0.878	0.830	0.884	1.000	
CpG5	0.915	0.827	0.916	0.909	1.000

B) *PAX8*

PAX8	***CpG1***	***CpG2***	***CpG3***	***CpG4***	***CpG5***
CpG1	1.000				
CpG2	0.931	1.000			
CpG3	0.956	0.946	1.000		
CpG4	0.951	0.918	0.944	1.000	
CpG5	0.961	0.930	0.942	0.950	1.000

C) *ZFYVE2*

ZFYVE28	***CpG1***	***CpG2***	***CpG3***	***CpG4***	***CpG5***
CpG1	1.000				
CpG2	0.787	1.000			
CpG3	0.850	0.830	1.000		
CpG4	0.671	0.663	0.658	1.000	
CpG5	0.705	0.709	0.748	0.683	1.000

D) *EXD3*

EXD3	***CpG1***	***CpG2***	***CpG3***	***CpG4***	***CpG5***
CpG1	1.000				
CpG2	0.952	1.000			
CpG3	0.961	0.951	1.000		
CpG4	0.859	0.866	0.897	1.000	
CpG5	0.932	0.942	0.960	0.941	1.000

E) *ZNF678*

ZNF678	***CpG1***	***CpG2***	***CpG3***	***CpG4***
CpG1	1.000			
CpG2	0.538	1.000		
CpG3	0.706	0.795	1.000	
CpG4	0.530	0.547	0.639	1.000

F) *PARD6G*

PARD6G	*CpG1*	*CpG2*	*CpG3*	*CpG4*	*CpG5*	*CpG6*
CpG1	1.000					
CpG2	0.868	1.000				
CpG3	0.829	0.909	1.000			
CpG4	0.827	0.858	0.872	1.000		
CpG5	0.765	0.814	0.864	0.833	1.000	
CpG6	0.754	0.752	0.753	0.795	0.782	1.000

G) *RBM46*

RBM46	*CpG1*	*CpG2*	*CpG3*	*CpG4*	*CpG5*	*CpG6*	*CpG7*	*CpG8*	*CpG9*	*CpG10*	*CpG11*	*CpG12*
CpG1	1.000											
CpG2	0.884	1.000										
CpG3	0.865	0.813	1.000									
CpG4	0.801	0.758	0.720	1.000								
CpG5	0.887	0.856	0.785	0.821	1.000							
CpG6	0.826	0.828	0.809	0.859	0.818	1.000						
CpG7	0.769	0.793	0.645	0.778	0.822	0.831	1.000					
CpG8	0.819	0.826	0.773	0.837	0.841	0.855	0.831	1.000				
CpG9	0.786	0.824	0.690	0.801	0.787	0.828	0.771	0.835	1.000			
CpG10	0.817	0.783	0.744	0.777	0.848	0.817	0.839	0.812	0.719	1.000		
CpG11	0.739	0.710	0.689	0.702	0.797	0.711	0.686	0.717	0.634	0.753	1.000	
CpG12	0.763	0.744	0.719	0.742	0.820	0.727	0.693	0.777	0.691	0.777	0.809	1.000

P-value< 0.0001 for all correlations shown in the above tables

Table 5.9: Correlation coefficients between mean methylation percentage across the loci assessed

Locus	*BOLA3*	*RBM46*	*PAX8*	*ZFYVE28*	*EXD3*	*PARD6G*	*ZNF678*
BOLA3	1.000						
RBM46	0.343**	1.000					
PAX8	0.023	0.094	1.000				
ZFYVE28	0.117	0.286*	0.236*	1.000			
EXD3	0.114	0.178	0.145	0.542**	1.000		
PARD6G	0.059	0.277**	0.244*	0.311**	0.144	1.000	
ZNF678	0.217*	0.352**	0.178*	0.197*	0.165	0.372**	1.000

N= 128 infants; Logit transformation of variables; *P< 0.05; **P< 0.005

5.1.3.4. Maternal biomarkers and infant DNA methylation

We investigated the impact of maternal blood biomarker status (i.e. one-carbon metabolites and cofactors) on DNA methylation patterns in their infants. To compress the information for the biomarkers PCA was performed. **Table 5.10** shows the coefficients by which each variable was multiplied to define the first principal component (i.e. the PC1 eigenvectors) of the biomarkers. Most biomarkers contributed positively to PC1, but the importance differed, from lowest for active B12 (0.074) to highest for homocysteine (1.525). The coefficients for folate, methionine and B2 (positive for EGRAC) had a negative sign. The PC1 of the biomarkers was significantly associated with season ($p<0.001$) and with methylation ($p=0.019$) (**Table 5.11**). However, when season and PC1 were included in the model together, neither was significant indicating possible correlation of PC1 with season. Interactions of DNA methylation with the PC1 of the biomarkers was also investigated, but none was significant. The PC1 of the biomarkers explained 22.6% of the overall variation in biomarkers.

To investigate whether there were common associations of each individual biomarker exposure with mean DNA methylation at MEs, all seven loci were tested jointly against each of the maternal biomarker variables separately, adjusted by infant sex and maternal BMI. Results with a p-value<0.05 are shown in Table 5.11.

Table 5.10: Principal component (PC1) coefficients for maternal biomarkers

Biomarker	1st principal component (PC1 $c1_j$)	PC1 $c1_j$/SD
Folate	-0.3441	-0.856
B2[1]	0.2864	1.116
Active B12	0.0314	0.074
Choline	0.2738	0.901
Betaine	0.1468	0.258
DMG	0.4794	0.950
Methionine	-0.1070	-0.586
SAM	0.1315	0.699
SAH	0.4132	0.967
Homocysteine	0.4874	1.525
PLP	0.0959	0.081
Cysteine	0.1611	1.105

For each individual PC1=$\sum_j c1_j.z_j$,where z_j= concentration for the j^{th} biomarker logarithmically transformed and $c1_j$ are the principal component coefficients for the j^{th} biomarker for PC1. This table tabulates $c1_j$ and $c1_j$/SD to standardise the coefficients.

[1]Coefficient for EGRAC (thus deficiency of B2 because higher EGRAC denotes greater deficiency)

Only active B12 and PLP (pyridoxal-5'-phosphate) were included in the model for B12 and B6 respectively

We previously observed in the group of non-pregnant women in The Gambia (research paper II, Section 4.1), that at times when B2 reaches the lowest concentrations, activity of the betaine-pathway is enhanced. We therefore investigated whether interactions existed between pathways, particularly between DMG concentration and lower/higher status for B2, but this was non-significant (results not shown).

Table 5.11: Associations between infant DNA methylation at seven metastable epialleles and environmental / maternal predictors

Biomarker/ Predictor	**ß coefficient** (95% CI)	**P-value**
Season	-0.323 (-0.517 - -0.1294)	0.001
Infant sex	-0.090 (-0.291-0.111)	0.382
Maternal BMI [years]	-0.044 (-0.076- -0.012)	0.007
Maternal age [kg/m^2]	-0.001 (-0.017-0.014)	0.850
Ethnicity mother	-0.117 (-0.514-0.279)	0.563
B2* [EGRAC coef][1]	-0.393 (-0.770- -0.017)	0.040
DMG:betaine ratio*	-0.213 (-0.366- -0.060)	0.006
SAM:SAH ratio*	0.319 (0.037- 0.601)	0.027
Cysteine* [μmol/l]	-0.840 (-1.555- -0.125)	0.021
PC1*	-0.067 (-0.130- -0.012)	0.019
PC2*	-0.052 (-0.017- 0.121)	0.139
PC3*	0.011 (-0.076- 0.099)	0.798

*Adjusted for infant sex and maternal BMI; Multilevel regression analysis; N=115; Biomarkers log-transformed

[1] Deficiency of B2 because higher EGRAC denotes greater deficiency);

DMG, dimethylglycine; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine; PC, principal component

5.1.4. DISCUSSION

Our study reveals that seasonality in rural Gambia affects methyl-group supply (as assessed by blood biomarkers of one-carbon metabolism) and DNA methylation in pregnant women and their infants, respectively. This demonstrates the importance of seasonality in this population setting and replicates our previously reported association between season of conception and DNA methylation at MEs. Contrary to our prior understanding, the rainy ('hungry') season was associated with higher methylation potential (as assessed by the SAM:SAH ratio and methyl-donors) and higher DNA methylation generally. Although a causal relationship between periconceptional maternal nutritional status and infant DNA methylation cannot be formally demonstrated with the current data, they appear to be associated and support the hypothesis of methyl-groups availability mediating the establishment of DNA methylation patterns.

5.1.4.1. Maternal blood biomarkers

All plasma biomarkers, except choline, B12 and cysteine, varied significantly between seasons and, with the exception of methionine and SAM, the direction of the seasonality is consistent with our findings in non-pregnant women (research paper II, Section 4.1). The SAM:SAH ratio is substantially higher (55.8%) in the dry season, suggesting a better methylation potential in the rainy season, possibly reflecting the higher availability of betaine, methionine, folate and B2 at this time of year. Conversely, the DMG:betaine ratio is lower (53.3%) during the dry season, suggesting higher activity of the betaine-dependent pathway, at a time when folate and B2 are lower. The extent of the differences between seasons differs by metabolite.

Regardless of the degree of matching of their seasonal patterns, the intercepts of most biomarkers are generally lower for pregnant than for non-pregnant women, as illustrated in Figure 5.2. This agrees with previous reports (118, 145, 169), and is

probably due to increased maternal and embryonic requirements and/or to physiological adaptations (e.g. hemodilution, increased glomerular filtration rate, and endocrinological changes (174)). These observed differences are unlikely to be explained by assay differences, since both sets of samples were analysed using the same protocols and equipment and at the same time. The effect of weeks of pregnancy since conception on biomarker levels (Table 5.3) agrees with those differences in intercepts shown in Figure 5.2, with negative association with the concentrations of B12, betaine, DMG, SAM, homocysteine and cysteine as weeks of gestation progress, whereas the association for folate is positive. Usually, pregnancy leads to a decline in folate status (118) over time. Whilst an increase in folate concentrations with pregnancy as seen in this study has been reported previously (145), it cannot be attributed here to the use of folic acid supplements, since blood collection was prior to antenatal registration and the administration of folic acid-containing supplements. Of note, the natural inter-seasonal difference in plasma folate levels observed in this population (5.48 nmol/l) is considerably higher than that found in Dutch women, supplemented or not with the currently recommended dose of 400 μg of synthetic folic acid during the periconceptional period (2.5 nmol/l) (166).

A strong negative correlation between homocysteine and folate concentration was observed (r=-0.374), as well as with B2 (r=0.195 correlation with EGRAC indicator of deficiency), which is consistent with other studies (175). Conversely, choline, betaine and DMG correlations with homocysteine are positive, as opposed to what has been found in The Dutch Antilles (145). It would be predicted *a priori* that folate and choline and betaine would both act similarly and decrease the plasma concentrations of homocysteine, by supplying methyl-groups. These positive correlations with the betaine-dependent pathway are difficult to interpret, as these substances are also involved in other metabolic pathways. Surprisingly, the SAM:SAH ratio seems independent of most of the biomarkers

and correlates only, negatively as would be expected, with choline (r=-0.225) and folate (r=-0.123). B2 does not seem to affect the SAM:SAH ratio despite being limiting in this setting (research paper II, Section 4.1).

5.1.4.2. Infant DNA methylation

In this study, we have expanded our previous research (79) on the season of conception effects on DNA methylation at MEs. Based on a different and larger cohort (born more than 10 years later than infants included in our earlier study (79)), we successfully replicated the observation of higher methylation at four previously reported MEs in infants conceived during the rainy season. We further present data on three newly identified MEs. The seasonal differences did not reach statistical significance for all seven loci (Table 5.5), but the clear and consistent pattern between loci suggests that seasonal factors play an important role in this setting. The consistency in response to season differs from the case reported by Tobi *et al*. on the Dutch Hunger Winter where the different loci varied in different direction in response to severe maternal malnutrition (75), and might be specific to MEs. A weaker effect within the current study for all four of the loci previously reported (79) can be observed and the reasons could be several. Firstly, this could be a result of a real secular variation in the seasonality (possibly nutritional but also of any other potential factor by which a seasonal effect on DNA methylation could be mediated, such as aflatoxin exposure (176, 177)). If the region is experiencing at least partly a nutritional transition, the seasonality could be subtler now than it used to be due to changes in accessibility, cash availability or in eating patterns, which could thus have reduced the effect of season of conception on differences in DNA methylation, compared to our previous study (79). Few data are available to try to elucidate this possibility, but a study on pregnant women from 1997 showed lower concentrations of B12 and homocysteine, and higher folate (97) compared to those observed in pregnant women

sampled in 2009-2010. Earlier data from 1978-80 indicated higher prevalence of folate deficiency but less of B2 (178, 179). However, these differences in biomarker concentration could also partially be due to technical differences in biomarker analysis. Secondly, it is possible that the Islamic Holy month of Ramadan played a role in the differences observed. During Ramadan, a period of fasting is observed from dawn until dusk and dietary practices are also know to vary with, for example, a greater consumption of sugary foods at the time of breaking the fast (180). In the current study, Ramadan occurred during August-September 2009, i.e. the rainy season period. However, the previous study (79) used samples from several years, of which 4 (namely 1991 to 1994) had Ramadan taking place during the dry season months. If Ramadan has any effect in terms of methyl-donors availability due to the qualitative and quantitative changes, it would have affected differently both studies (i.e. if Ramadan was favourable for DNA methylation it could have enlarged the differences and *vice versa).*

From our departure hypothesis of methyl-donor availability explaining the seasonal effect, we would expect that all MEs responded similarly to season and maternal biomarkers. From our data (Table 5.6), methylation at some individual loci appears more sensitive to environmental influences. The mean methylation of the two loci which showed statistically significant associations with season (*PAX8* and *RBM46*) were not correlated. Similarly, the poor pairwise correlation coefficients between mean DNA methylation across all seven MEs studied would suggest that methylation is not strongly driven by any individual external factor such as, for example methyl-donors availability. Thus, strong regulatory factors or random processes could exist, leading to variable responses to environmental factors for each locus. However, interactions between methylation and season, and differences in response between loci, were statistically tested. These were found not to be significant, indicating there is no difference in the response to season between the different MEs under study. Therefore it cannot be

concluded that individual loci respond differently to season. The power to detect such interactions was, however, relatively weak.

Interestingly, BMI was the maternal predictor of infant DNA methylation level with strongest statistical significance, above any of the maternal biomarkers under study. This indicates for the first time a (negative) association of a maternal nutritional condition in early pregnancy with methylation at MEs and supports the importance of maternal nutrition in the establishment of foetal DNA methylation patterns. This is contrary to findings from Argentina where maternal BMI appeared to be positively correlated to *PPARGC1A* promoter methylation in DNA from umbilical cord blood (181). However, *PPARGC1A* is not known to be an ME and so, not directly comparable. Furthermore, our findings are in line with our knowledge of the weight loss that takes place during the rainy season (126).

The specific loci assessed in this study were chosen solely on their characteristic of being MEs (and thus sensitive to environmental changes during early pregnancy and with no intra-tissue variability). Yet, proteins encoded by genes described here are known to have metabolic effects and even to be related to disease (e.g. *PAX8* is involved in hypothyroidism (182). However, whether higher or lower DNA methylation at those genes is more desirable phenotypically was not considered within this proof of principal study and would form the focus of future work. The focus of this study was in the biological determinants of such methylation.

5.1.4.3. Maternal biomarker relationship with infant DNA methylation

The ultimate aim of this study was assessing, in humans, whether observed effects of season of conception on infant DNA methylation are related to one-carbon metabolite status of their mothers during early pregnancy. The different MEs under study (and their

CpGs) could be differently regulated by or protected from environmental cues. However, a reasonable first assumption would be that biomarkers had a common effect on MEs. Interestingly, when methylation across all loci combined was assessed, we did see an association with the SAM:SAH ratio (indicating the methylation potential) and the DMG:betaine ratio, as well as with cysteine and B2. Further, the PCA showed that one-carbon metabolism biomarker PC1 was significantly associated with methylation. B12 was not indicative of lower infant DNA methylation at MEs in our study, in contrast to a recent report on global DNA methylation (78). SAM, which has a central role as substrate the provision of methyl-groups (Figure 4.1) and its product, SAH, were not associated with DNA methylation individually. This substrate and product role in methyltransferase reactions is why the ratio between SAM and SAH has been proposed as metabolic indicator of cellular methylation status, representing the activity of both pathways simultaneously in an intrinsically complex metabolic process. It is thought that SAH regulates transmethylation and it has been shown that chronic elevation of SAH, secondary to the homocysteine-mediated reversal of the SAH hydrolase reaction can reduce methylation of DNA (183). The severity of B2 deficiency seems to limit folate use in this population as well as DNA methylation, which in this population seems more associated to the betaine pathway (DMG:betaine ratio).

The degree of inter-individual variation in prenatal DNA methylation and the factors that regulate and affect the establishment of methylation patterns in early development are poorly understood to date and evidence is scarce. Our findings support the hypothesis that methyl-donor supply could drive the DNA methylation in early development. However, both the effect size and the significance are weak and could be due to confounding in this observational study. Therefore, the data have to be interpreted with caution, and require replication. For example, the association observed between the biomarker PC1 and the MEs might be due to PC1 picking up on the seasonality as the

strongest common factor between biomarkers. In fact, PC1 seems to be correlated with season ($r = 0.569$, p-value< 0.0001), and both become non-significant if both are included in the regression model.

In addition, PC1 coefficients for folate, B2 and methionine are negative, suggesting that methyl-group availability could not be the main explanation. In this population, women do not smoke and are not exposed to environmental arsenic, which are known factors to affect DNA methylation. However, women are highly exposed to cooking smoke (184, 185), although exposure does not seem to differ by season (pilot data, not shown). The effect of one-carbon metabolism could respond to: 1) methyl-group availability for DNA methylation, or 2) deliberate programming in response to inappropriate environment, as assessed by methyl-group deficit (and/or other environmental seasonal effects). Due the critical role of the DNA methylation pattern establishment in early development and the redundancy of methyl-groups supply paths and variety of sources to ensure supply, the existence of tightly regulation in these epigenetic processes is reasonable.

Additionally, the variation in nutritional status between seasons observed in this population, although clearly significant and of importance (research paper II, Section 4.1), is limited in comparison for example with the dramatic dietary intake reduction during a state of severe malnutrition as was the case during the Dutch Hunger Winter (74). This might explain why the strength of the association of maternal nutritional status with MEs is limited in our current data. Furthermore, it is possible that maternal one-carbon metabolites could have a stronger effect on global DNA methylation (currently under investigation) or on imprinted genes.

5.1.4.4. Limitations of the study

Our hypothesis was that DNA methylation at MEs is established close to conception. Adjustment of the plasma biomarker data by back-extrapolation to the date of conception using data from a separate group of women seemed therefore necessary. We acknowledge that there is variability in the suitability of these biomarkers for back-extrapolation from the group of non-pregnant women. Particularly, our measures of SAH and DMG were shown to have limited reliability and stability, either due to intra-individual fluctuations or to measurement issues (research paper II, Section 4.1), which can introduce error. Pregnancy could also distort the seasonal patterns observed in the biomarkers as discussed above. The current study has an increased sample size compared to the previous report (79), however it remains a relatively small study, which might reduce the ability of our study to detect significant effects. Still, this represents the largest study of this kind in humans. In addition, due to the cyclic and dynamic nature of one-carbon-metabolism, the redundancy of pathways and complex regulation, a static single measurement may be insufficient to give us a real picture of the methylation capacity. Incorporation of serial measurements as well as enzymatic activity could help in the interpretation.

Finally, key factors to look at in this or in future studies are maternal and infant genotype in this population relative to SNPs of interest. We have not determined the maternal or infant genotype for SNIPs in 5,10-methylenetetrahydrofolate reductase (MTHFR) and other choline and one-carbon metabolism genes, which affect their blood levels and thus could perhaps explain some of the findings (175). Frequency information on such SNPs is sparse in our study population has been sparsely studied in our population or other comparable African populations.

5.1.4.5. Conclusions

This study replicates and expands on the findings from our previous study (79) (research paper II, Section 4.1) that the Gambian rainy season, traditionally considered harder and more nutritionally challenging, is associated with higher levels of methyl-group supply and DNA methylation. Further, this study provides evidence in humans that maternal nutritional status (including BMI and plasma one-carbon metabolites) predicts infant DNA methylation at MEs in early development. The replication and the consistency of the seasonal effect on DNA methylation indicate the important role of the environment. In order to explore whether causal relationships exist, supplementation trials will be needed.

ACKNOWLEDGEMENTS

We thank the women of West Kiang for their kind participation, and all the staff at MRC Keneba. Also thanks to Dr. Dorothea Nitsch (LSHTM) and Dr. Chris Bates and Lorna Cox (MRC HNR, Cambridge) for their valuable advice and input to the study and to Jannette King (UBC) for the laboratory analyses.

AUTHOR CONTRIBUTIONS

Conceived and designed the study: PDS, SEM, SEC, RAW, AJCF, AMP & BJH. Conducted the research: PDS, SEM & BJH. Conducted laboratory analyses: RD, SMI and RAW. Analysed the data: PDS & AJCF. Wrote the paper: PDS drafted the manuscript and all other authors critically reviewed and approved it. No conflict of interests was declared.

5.2.ADDITIONAL METHODS

5.2.1.Study population

This main group study was conducted in tandem to the ENID trial (see Section 3.2.2): the round of meetings in each village to explain to the local community the full details of the studies and to seek their approval, the recruitment waves and the monthly LMP questionnaire rounds were done in tandem. Participants were given the option to participate in both, one or neither of the studies. Owing to a delay with the manufacture of the prenatal supplements, the ENID trial only started in full late-April 2010. Since the study was designed to ensure that women were not forced to disclose their pregnancy early, all blood sampling for the purpose of the current study was performed prior to pregnancy confirmation and thus prior to supplementation as part of ENID. Once the women were identified as positive by urinary dipstick (Appendices III (D) and XII), they were invited to MRC Keneba for confirmation of pregnancy, gestational age assessment by ultrasound examination and a full antenatal examination. Focused antenatal care was offered to the enrolled women during their pregnancy through the routine trekking service, with women brought into MRC Keneba at 35 weeks gestational age for a final pre-delivery check.

5.2.2.Maternal blood biomarkers

Once every four weeks, women enrolled into the Methyl-Donors and Epigenetics (MDEG) study were invited to a central location in their village of residence for an LMP questionnaire by a fieldworker resident in or near their village. Women were requested to fast prior to the visit. A trained nurse collected blood samples from all women who reported a missed menses, but without an 'exempting reason' (see Figure 3.3). All village women attended throughout the morning and therefore, the time from collection to

storage, after transport to MRC Keneba and processing varied between samples collected at the beginning or at the end of the morning. Women who did not attend the call were visited in their homes during the following days by the posted fieldworkers. Only when the participants were known to be eligible (as assessed later by their conception date and their gestational age at blood sample collection), their samples were sent for analysis to the collaborating laboratories in Canada or the UK. The remaining samples were excluded. Methods for the processing and analysis of blood samples were the same as the indicator group (Section 4.1.2.3) and the samples were treated identically (i.e. using the same protocols and techniques for processing and analysing, stored and shipped together, and analysed within the same timeframe).

5.2.3. Infant DNA methylation

DNA methylation patterns are the key outcome in this study. The choice of MEs was based on their particular nature. MEs are loci at which systemic interindividual epigenetic variation occurs independently of genotype. Animal models showing an effect of maternal diet in fetal programming were mostly based in this type of loci (69). As discussed in Section 2.3.4. previous research (104) was successful in identifying loci in humans that fulfilled the established definition to be an ME:

(i) stochastic establishment (resulting in dramatic and systemic interindividual variation in locus-specific epigenetic regulation, with discordance in monozygotic twins);

(ii) established early in development (around the time of conception);

(iii) no tissue-specific methylation (established in the early embryo and then maintained in all germ-layer lineages); and

(iv) observed susceptibility to nutrition and other environmental cues (i.e. responding to season of conception) (79).

These characteristics make MEs good candidate loci to test epigenetic mechanisms in the developmental origins hypothesis and where used as criteria to identify candidate MEs from a MSAM screen of the genome, after a round of eliminatory steps to test: i) systemic interindividual variation through correlation across tissues in autopsy samples; ii) substantial discordance in monozygotic twins; and iii) persistent epigenetic effects of an environmental exposure (i.e. Gambian seasonality) (79).

Since the publication of the first five MEs (Figure 2.7) by Waterland *et al*. (79), further analyses have been conducted on liver, kidney and brain tissues from additional donors (n=17) for further assessment of inter-tissue correlation of methylation and formalise the definition of candidate MEs. SLITRK1 was exceptional in that methylation in brain did not correlate with that in liver and kidney (79). Therefore this gene was dropped according to a stricter criterion on inter-tissue correlation and not investigated in the current larger sample cohort, however, the other four previously described loci were tested in this study. Additionally, three newly identified (and as yet unpublished) MEs were also included in the present sample set (*PARD6G*, *RBM46* and *ZNF678*).

As yet it is not known whether variation at these MEs lead to any functional consequences in phenotype. *PAX8* is known to be involved in hypothyroidism (186) but the role of other loci in their metabolic pathways is unclear.

The role of methyl-donor supply in modifying DNA methylation at MEs and phenotype/disease outcomes thereof has been shown in animal models (69). The goal in this exploratory work was to assess as a first step whether – as in animal models – methyldonor supply affects methylation patterns at MEs in humans. The aim was not to identify implications of MEs, which will be considered in further analyses beyond of the scope of this proof-of-principle thesis (as discussed in Section 6.3.).

Infants were brought to MRC Keneba at 3-9 months of age for blood collection and anthropometric measurements (weight, length, mid-upper arm circumference (MUAC)),

all performed using regularly validated equipment and following standard procedures. Their personal welfare card, which tracks the infant's progress from birth as a part of the national Maternal and Child Health Programme, was copied and relevant information extracted (i.e. birthweight and follow up measurements of weight and length, when available). The information included in the welfare card was, however, incomplete and was therefore not used.

5.2.4.Statistical analysis

Study population

Based on previous information of birth rates, 100 individuals per season was an achievable number. The initial power calculation for these 100 individuals per group is shown in **Table 5.12**. Unfortunately, it was not possible to reach this target, and the reasons for this are discussed in Section 5.3.1. Approximately 83 women and 65 infants per group (depending on the biomarker and the loci considered respectively) had their biomarker concentrations and DNA methylation analysed. The different potential sample sizes are compiled in Table 5.12 with the difference detection allowed for each. To test the main hypothesis (differences in DNA methylation between infants born in the rainy and dry season), the current sample of 60-65 infants allowed us to detect a difference of 0.57-0.59 standard deviations, with power 90% at 5% significance.

The statistic methods for the biomarkers and DNA methylation are described in the Methods section of the paper in Section 5.1.2.4.

Table 5.12: Difference detection (in standard deviations) for different sample sizes for a two-group comparison, with any given reliability

	$\|\Delta\mu\|/\sigma$				
N (per group)	**Reliability**				
	1	**0.8**	**0.6**	**0.4**	**0.2**
60	0.59	0.66	0.76	0.94	1.32
65	0.57	0.64	0.73	0.90	1.27
80	0.51	0.57	0.66	0.81	1.15
85	0.50	0.56	0.64	0.79	1.11
100	0.46	0.51	0.59	0.72	1.03

$|\Delta\mu|$: difference in means; σ standard deviation of the variation between individuals

N: sample size

$|\Delta\mu|/\sigma = \Phi^{-1}[1-\alpha/2] + \Phi^{-1}[\beta]/\sqrt{(N \times \text{reliability}/2)}$

Φ^{-1} inverse of the standard normal distribution function

α (probability of a Type I error occurring or significance)

β probability of a Type II error occurring or power)

Maternal blood biomarkers

Full details of the analysis performed on maternal blood biomarkers are given in Section 5.1.2.4.

Infant DNA methylation

Full details of the analysis performed on infant DNA methylation are given in Section 5.1.2.4.

Maternal biomarkers and infant DNA methylation

Full details of the analysis performed on the influence of maternal biomarkers and infant DNA methylation are given in Section 5.1.2.4.

5.3.ADDITIONAL RESULTS

5.3.1.Study population

Two waves of recruitment took place. Most women were recruited in the first wave (July 2009), when all women 18-45 years of age throughout West Kiang were approached for consent. The second wave (February 2010) aimed at recruiting new women moving into West Kiang, or women who were not eligible for recruitment during the first wave (i.e. pregnant at the time, below 18 years or temporarily out of the region). **Table 5.13** compiles the information on the recruitment data, based on the inclusion / exclusion criteria described in Section 5.1.2.1. If women moved after they had been identified as pregnant and were eligible for either of the pre-defined seasons and had a gestational age at blood sample collection below 16 weeks, they stayed in the study and their infants were invited to come to MRC Keneba after they were born.

As discussed in Section 5.2.4., the number of mother-infant pairs participating in the study was lower than initially planned. This was mostly due to the difficulty in the early identification of pregnancies, and was regardless of the prevention measures taken to avoid this (i.e. sensitisation meetings by the midwife and confidentiality assurance). Despite efforts to follow women up each month, sometimes missed menses passed undetected because women were temporarily travelling and thus were not seen. Further, it was sometimes reported that menses were being seen even when the women were several months pregnant. This could be due to the social issues relating to early pregnancy disclosure (187), with some women even withdrawing from the study once they suspected they were pregnant. Of all the children participating in MDEG and being conceived during the dry season, 52 (40.6%) were co-enrolled in the ENID Trial.

Table 5.13: Recruitment summary

	Women approached[2]	Pregnant at consent (excluded)	Eligible women for consent	Consented women	Participants leaving the study[3]
1st wave of recruitment[1]	2233	143	2090	1630	295
2nd wave of recruitment[1]	813	40	773	417	40
Total	3043	183	2863	2047	335

[1]The 1st wave of recruitment started in July 2009 and the second in February 2010

[2]Women 18-45 years of age in West Kiang

[3]Up to July 2010: this includes women withdrawn by the study (i.e. mostly found to be menopausal, but also moving out from the region or long-term travel) as well as self-withdrawal (main causes being husband disapproval and, occasionally, specific factors such as divorce or hysterectomy).

5.3.2. Maternal blood biomarkers

Table 5.14 shows the overall geometric means of the biomarkers for the pregnant women in the main group, including levels of sufficiency from reference values.

5.3.3. Infant DNA methylation

The results are described in the manuscript presented in Section 5.1.3.3. The boxplot graph in **Figure 5.3** shows graphically the correlations within each gene among its CpG sites (Table 5.9). It can be observed that there were few 'outlying' correlation values. Mean values of DNA methylation were very high for *PAX8*, *EXD3*, *BOLA3* and *PARD6G* (>80.0%). The lowest level of methylation was seen for *ZNF678* (60.38%).

Table 5.14: Mean blood biomarker concentration in rural Gambian women

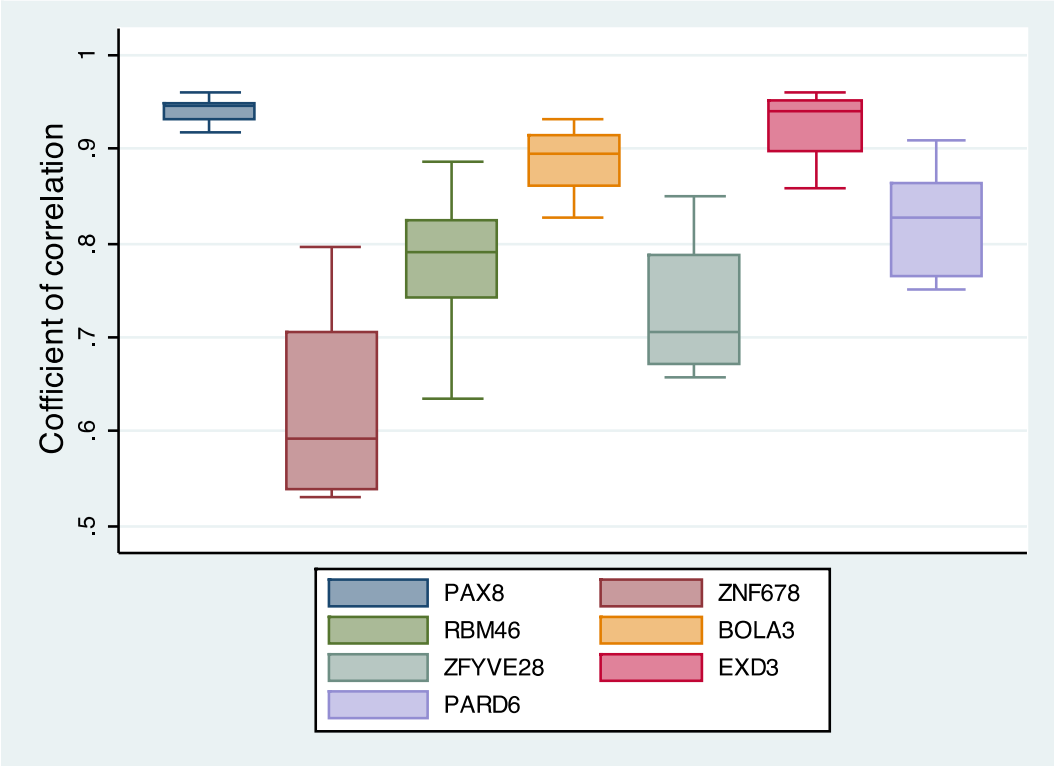
Biomarker	Geometric mean	95%CI	% of women below cut-off values
Folate [nmol/l]	16.45	15.50 - 17.46	12.19%
B12 [pmol/l]	318.43	298.92 - 339.23	17.68%
Active B12 [pmol/l]	69.75	64.74 - 75.15	7.32%
B2 deficiency$_1$ [EGRAC coef]	2.17	2.09- 2.25	98.71%
Choline [μmol/l]	6.83	6.52 - 7.16	90.85%
Betaine [μmol/l]	19.56	18.05 - 21.20	-
DMG [μmol/l]	1.76	1.62 - 1.92	-
Methionine [μmol/l]	23.50	22.79 - 24.24	-
SAM [nmol/l]	78.63	76.57 - 80.75	-
SAH [nmol/l]	11.94	11.37 - 12.53	-
SAM:SAH	6.59	6.29 - 6.89	-
Homocysteine [μmol/l]	7.10	6.77 - 7.43	0.61%
Cysteine [μmol/l]	197.20	192.96 - 201.53	-
PL [nmol/l]	4.37	4.14 - 4.61	-
PLP [nmol/l]	26.56	24.64 - 28.62	31.7%
PA [nmol/l]	9.47	8.86 - 10.12	-

N ranging between 163-164 depending on the biomarker

[1]B2 def: B2 deficiency as assessed by EGRAC; DMG, dimethylglycine; Hcy, homocysteine; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine.

(-) is indicated where not suitable cut-off could be identified

Figure 5.3: Boxplot of intra-gene correlations between CpG sites



Logit transformation of variables for the correlations

5.3.4. Maternal blood biomarkers and infant DNA methylation

The initial assumption was that MEs, by definition, would act similarly in response to the shortage / availability of methyl-groups, and this was investigated in the research paper III (Section 5.1.3.4), combining the average loci methylation of the seven MEs. Nevertheless, there is a possibility that different CpGs / loci might have different levels of protection / susceptibility to stochastic changes to regulation of methylation. Therefore, the individual associations of each biomarker exposure with average DNA methylation at each ME was also assessed by multilevel regression, including all the individual CpG sites for each ME. Results with a p-value<0.05 are shown in **Table 5.15**. It was observed that only cysteine was significantly (negatively) associated with methylation across several genes, namely BOLA3, RBM46 and ZNF678. Other biomarkers were only associated with methylation of one single gene (mostly RBM46, which is the locus with more potential for

variation since the number of CpGs (12) is much higher than for the other loci under study, varying from 4 to 6). At only one ME (*RBM46*) DNA methylation was associated with SAH and no ME was corrrelated with SAM or the SAM:SAH ratio. The significances were weak and none was below a p-value of 0.0005 (following Bonferroni correction). Furthermore, when the interactions between methylation at the different loci and the individual biomarkers were assessed, the differences between loci resulted non-statistically significant, suggesting that the MEs under analysis are not likely to exhibit differential associations with biomarkers.

Assessing the effect of biomarker interactions on methylation outcome in this set-up is difficult due to the limited sample size and the multiple testing that this involves. However, I did some explorative analysis to see whether any very strong interaction were apparent. The interactions between each pair of biomarkers on the DNA methylation at each locus were thus assessed. The regression analyses were all non-significant at the $p<0.0005$ level (following Bonferroni correction) and the data are not presented here. Bonferroni correction is known to be very restrictive; however, even with a 10-fold higher significance cut-off (p-value<0.005), only the following interactions were significant: active B12-homocysteine ($\beta=0.565$; $p= 0.004$), B2-SAH ($\beta=-0.852$; $p= 0.001$), choline-cysteine ($\beta=2.360$; $p= 0.004$), and DMG-cysteine ($\beta=1.620$; $p= 0.004$), each one in a different gene. These data do not thus show consistency or any clear trend that can explain any pattern in inter-metabolite interactions. When all loci were assessed with multi-level analysis, only the interaction between B2 and SAH was below p-value<0.005 ($\beta=-1.353$, $p=0.003$).

Table 5.15: Associations between mean infant DNA methylation at metastable epialleles and maternal biomarkers

Biomarker/ Predictor	Methylation outcome	N	ß coefficient (95% CI)	P-value
B2 [EGRAC coef]	*RBM46*	113	-0.273 (-0.483 - -0.063)	0.011
Betaine [μmol/l]	*ZFYVE28*	115	0.079 (0.012 - 0.145)	0.021
DMG [μmol/l]	*RBM46*	108	-0.164 (-0.257 - -0.036)	0.010
Methionine [μmol/l]	*PARD6G*	110	0.355 (0.034 - 0.676)	0.031
SAH [nmol/l]	*RBM46*	108	-0.145 (-0.271 - -0.019)	0.025
Homocysteine [μmol/l]	*RBM46*	114	-0.287 (-0.456 - -0.117)	0.001
Cysteine [μmol/l]	*BOLA3*	115	-0.519 (-0.952 - -0.085)	0.019
	RBM46	115	-0.476 (-0.878 - -0.074)	0.021
	ZNF678	114	-0.229 (-0.414 - -0.045)	0.015
B6 [nmol/l]	*BOLA3*	115	-0.173 (-0.323 - -0.022)	0.025

Adjusted for infant sex and maternal BMI; Biomarkers log-transformed

Multilevel regression analysis

Cut-off of significance with Bonforroni correction = 0.0005

5.4.ADDITIONAL DISCUSSION

5.4.1.Maternal blood biomarkers

A major limitation of the maternal blood biomarker assessment was the difficulty in collection of a close-to conception blood sample. The low numbers of participants recruited early in pregnancy led to a forced expansion of the cut-off initially set (up to 12 weeks), up to 16 weeks. This had two implications to the results: firstly, the women were in a more advanced stage of gestation than planned in the original study design which may be relevant given that certain aspects of physiology (including one-carbon metabolism) are notably different from non-pregnant women (169). Secondly, this led to a wider time variation in the window of actual sample collection and hence biomarker assessment. For example, within the rainy season months of July to September, a time gap of almost 6 months may exists between a woman conceiving on 1st July and bled one week post conception to a woman conceiving on 30th September and bled at 16 weeks post conception, with implications for the assessment of seasonal differences. Only 42 women for the rainy season and 27 women for the dry season had their blood sample collected within the established timeframe (between July and September and between February and April respectively). Hence, the correction (by back-extrapolation using indicator group data) is appropriate and in fact essential. Correction of the recorded values is the best option for the analysis to be meaningful, without distortion induced by seasonality. Not considering the nutritional status at the right season of conception would jeopardise the whole objective of the study. Unfortunately, such extrapolation has its own caveats and introduces a different type of error, as discussed in 5.1., particularly for those biomarkers where reliability and stability is poor (e.g. SAH or DMG, see Figure 4.2). Figure 5.2 showed that the appropriateness for this back-extrapolation could differ widely between biomarkers, from those showing very similar patterns between main group and indicator group (folate, B12, active B12, choline, betaine, DMG/betaine ratio and SAM) to those

where patterns seem to differ (B2, DMG, methionine, SAH, SAM:SAH ratio, homocysteine, cysteine and B6). Both SAH and DMG are the least informative biomarkers, as assessed by their reliability and stability, and therefore they need to be considered with special care. This can lead to over- or under-estimation of the seasonal differences and of the biomarker concentrations at conception, affecting downstream analyses of these as exposure. Although the sample size was very limited, a sensitivity test was done by repeating the analyses, using only the women whose sample was drawn within the defined season windows. It was observed that most biomarkers showed the same directionality of differences as the initial back-extrapolated analyses when values from women at each season were compared. The exceptions were cysteine, choline and B2, all of which were biomarkers with divergence of seasonal pattern from the indicator group. This indicates the need to interpret these biomarkers with caution. However, these divergent differences between seasons was only significant for cysteine (p-value = 0.0011). Thus, the use of back-extrapolated data was deemed an appropriate means to employ in the overall analyses.

Adjustment was applied for those variables which varied with time of processing, namely DMG and SAH (see Section 5.1) for all subsequent analyses. The ideal time from sample collection to storage at -80ºC, particularly for SAM and SAH stability and avoidance of interconversion of SAM to SAH, is two hours, according to experience from the Innis laboratory (data not published). Due to geographical area of recruitment and distances, as well as field conditions (quality of roads and population spread), there were more issues with longer time of processing in the main group compared to the indicator group. Thus, the adjustment for the main group becomes more important.

Regarding biomarker status, concentrations of B2 and choline seem particularly low in this population, as compared to recommended cut-offs (120, 122). Bates *et al.* (95) described a widespread B2 deficient status in pregnant and lactating women in The

Gambia 30 years ago, and a higher prevalence of deficiency (higher EGRAC) during the peak of the dry compared to the peak of the rainy season (95). B2 deficiency can impair the folate dependent pathway within one-carbon metabolism, potentially leading to a switch to the folate-independent (betaine-dependent) pathway (111), which might be impaired by the low choline. The apparent choline deficiency is probably compensated for by betaine in the methyl-group donation, since the women in the current study do not show elevated plasma homocysteine (>15 μmol/l) and B-vitamins deficiencies do often result in disturbance of one-carbon metabolism, reflected in a rise in homocysteine (111).

Interestingly, when looking at the relative plasma concentrations of choline, betaine and DMG in non-pregnant women in this population (1:4.9:0.4 (Section 4.1)), these were similar to the proportions reported in Canadian non-pregnant women (1:4.5:0.3) ((169)). However, for the pregnant women the proportions differed substantially (1:2.5:0.3) in relation to Canadian pregnant women (1:0.9:0.1), where choline becomes the major plasma metabolite. This is probably linked to the fact that in the Canadian women, choline increased in pregnancy whilst in the Gambian women it decreased. In contrast with the steady increase of choline that has been shown in pregnancy in other populations (145, 169) due to oestrogen-induced upregulation of endogenous synthesis, the values for choline for these women were lower for pregnant (the geometric means in both seasonal groups were below 7 μmol/l) than for non-pregnant women (geometric mean 8.2 μmol/l) (Section 4.1.3.3). The fact that these samples were collected in earlier stages of pregnancy is likely to explain this finding. It is possible, however, that in a more nutritionally challenged population, homeostatic systems protect one-carbon metabolism to a certain extent. The demand for choline is high during pregnancy due to its critical role in embryonic life, not only for DNA methylation regulated gene expression, but also for the appropriate development of the brain and the nervous system (choline is involved in building cell membranes and in

acetylcholine for cholinergic transmission and transmembrane signalling ((188)). Given the very limited choline intake from diet in West Kiang women (Section 4.1.), it is plausible that the endogenous synthesis is not enough and thus, a maternal plasma decrease is observed due to placental transfer to the foetus.

The plasma biomarker levels studied here are static measurements at a single point in time. As such, these levels mainly reflect recent intakes (118), more than the mid/long-term status, with the exception of the RBC B2 EGRAC. In order to assess more accurately one-carbon metabolism status in a dynamic fashion, the metabolites should be looked at together with the activity of the enzymes that control the processes. Disruptions to any of the pathways could considerably distort the scene, and the dynamic relationships would give us information not only on the pools but also the flux. One-carbon metabolism is linked to the larger protein pool mainly via methionine, which could suggest that sufficient methionine availability exists. One-carbon metabolism also related involves other amino acids such as glycine or threonine, which intervene in the folate-dependent pathway and can affect the SAM:SAH ratio balance (189). SAM and SAH plasma concentrations are frequently used as markers of global DNA methylation potential (183). It is believed that SAH could be more critical than SAM in the regulation of methylation. This makes sense, since DNA methyltransferase enzymes (DNMT) bind SAH with higher affinity than SAM and, therefore, they are subject to potent product inhibition by SAH (190). Higher levels of homocysteine that do not get remethylated via betaine- or folate-dependent pathways or, alternatively, transsulfuration, due to deficiencies, could reverse to SAH and slow methylation processes (111). BMI has shown to be a strong determinant of SAM and SAH (175), but did not show any effect in our study on any of the biomarkers assessed, other than B12 (total and active).

The potential influence of the Ramadan has been discussed in Section 5.1.4.2. in the context of differences in DNA methylation between this study and the past study on

DNA methylation at MEs (79). With respect to the biomarkers, the Ramadan taking place during the peak of the rainy season 2009 is also worth a mention, since it could have affected those plasma biomarkers sensitive to recent dietary changes. Pregnant women are exempt from fasting if they are worrying about their own health, or the health of their foetus (191). Information was not collected on women who fasted or not to control for this potential confounder. However, in the past, it has been observed that 90% of pregnant Gambian women would fast (180) in this particular setting. The tendency is to fast as far as pregnant women feel physically well, at least during the first 6 months of pregnancy. This is because Ramadan whilst a religious event, is also a social one. Pregnant women who choose not to fast during those days, would have to fast later in the year to compensate, once the pregnancy is over, as well as, paying back a compensation with alms (gift in the form of money or food to the poor) (191). It is certainly more difficult to fast outside the Ramadan period, when all the rest of the family / community is not fasting (192). As discussed for the indicator group, some evidence supports that no major quantitative changes take place during the Ramadan (i.e. similar energy intake). However, qualitative changes and thus micronutrient contents could be more affected (150) and because Ramadan occurred in the peak of the rainy season, this fact might have faded out or increased seasonal differences.

5.4.2. Infant DNA methylation

Differences (reduced seasonal effects on methylation) compared to previously published data (79) could be due to methodological issues, namely the use of slightly different conception timeframes considered for this study (i.e. July-September *vs.* August-September in the previous study and February-April *vs.* March-May previously used). Thus, the identical analyses were repeated with the data from the present study, but using only the those samples from infants with comparable conception dates, that is, conceived between August and September and March and April (for the previous study, the time

window was March to May, but May samples were not available for the present study). The results obtained were not dramatically different from the results presented in Section 5.1.3.3 in terms of the statistical significance of the seasonal differences, and there were only two MEs statistically significant, but these were two different MEs (see Table 5.16). Using the alternative cut-offs, again methylation was consistently higher during the rainy season. The total level of gene DNA methylation was generally higher in the rainy season, and mostly lower in the dry season, but these differences were slight and did not generally reach the methylation levels of the previous study. Data from May, had they existed for the present study, might have led to more important changes, but this cannot be ascertained. The lack of comparability in the results obtained suggests that the differences seen between the earlier and the current data are more likely to be related to other factors (e.g. secular differences since 1991) rather than to protocol differences.

Table 5.16: Mean DNA methylation at metastable epialleles and season of conception according to cut-offs set in Waterland, PLoS 2010.

Locus (N CpG site)	**Rainy season mean methylation** [%] (95% CI)	**Dry season mean methylation** [%] (95% CI)	**Between season difference** Δ%[a]	**P-value**
BOLA3 (5)	51.93 (47.23-56.62)	51.33 (46.37-56.29)	0.6	0.8604
PAX8 (5)	77.54 (72.04-83.04)	74.24 (68.25-80.22)	3.3	0.4172
RBM46 (12)	70.45 (67.37-73.53)	67.79 (64.71-70.87)	2.66	0.2223
ZFYVE28 (5)	31.61 (27.11-36.10)	23.57 (19.48-27.65)	8.04	0.0089*
EXD3 (5)	81.42 (78.91-83.92)	79.09 (76.44-81.73)	2.32	0.2028
PARD6G (6)	55.06 (50.54-59.58)	47.08 (41.72-52.45)	7.98	0.0251*
ZNF678 (4)	32.03 (28.75-35.32)	28.62 (26.19-31.06)	3.41	0.0909

[a] Mean absolute difference in DNA methylation between dry and rainy season; * p-value < 0.05

N = 45 for rainy season and n=57 for dry season unless otherwise indicated

As shown in Figure 2.7, the previous study found a higher average level of DNA methylation levels for three of the four loci also assessed in this study (i.e. 50-51% for *BOLA3* in our study as opposed to 55-61% in the past study; 71-76% for PAX8 as opposed to 70-80%; and 25-29% for *ZFYVE28*, as opposed to 27-38%) (79) and similar for EXD3. However, the broad ranges of methylation at these three loci is very comparable between the studies, again pointing towards the reproducibility of the results from the earlier work in this population.

Therefore, it is unlikely that differences with the previous study are due to sample size. It is not unrealistic to think that a seasonality effect might have faded out over time or that any other specific environmental factor relevant to the period of study could have played a role. As previously discussed, Ramadan affected the last month of the rainy season peak window. It has been shown in a pilot study in Turkey that effect of Ramadan fasting during the second trimester does not seem to have a significant effect on maternal oxidative stress, foetal development or birth weight (193). However, it does not exclude that Ramadan can induce these or other changes in early pregnancy, which could be more sensitive period in terms of DNA methylation, as indicated in the Dutch Hunger Winter study (74). In fact, Almond *et al*. (194) have recently shown that Ramadan fasting during the first month could be associated with lower birth weights. Then, hypothetically, if this fasting during the rainy season did affect negatively the birth weight and potentially DNA methylation, this could worsen the rainy season, weakening the inter-season differences.

Tobi *et. al* (75) found sex-specific differences in levels of DNA methylation in humans (75, 195), an observation which is important to consider in DNA methylation studies. Further, recent DOHaD research has suggested that males could be more ready than females to trade off visceral development *in utero* to protect somatic and brain growth, as well as being more responsive to their mother's current diet than females, who are more responsive to their mother's life-time nutrition and metabolism (196). Here, we

found sex-specific DNA methylation for three genes; higher levels for males at EXD3 (81.5% *vs.* 78.5%), and higher for females at RBM46 and *ZNF678* (72.5% *vs.* 67.5% and 33.2% *vs.* 28.5% respectively). However, sex did not seem to interact with the seasonal effect (i.e. the sex-specific differences did not vary by season). Recent evidence also suggests that DNA methylation can be influenced by various other factors, which were investigated. However, none of the genes seemed to be affected by total length of gestation (78, 111), supporting that MEs do not change their methylation status over time. Maternal age at pregnancy (197) and self-reported ethnicity (198) had no effect on methylation status. Ethnicity in West Kiang is predominantly Mandinka, although Fula and Jola's are also present, as shown in Figure 2.4. For the infants we asked specifically for the mother and father's ethnicity separately, but it must be noted that in the case of mixed couples, the ethnicity reported would be that of the father. Finally, in this study data was not collected on several other covariates, that would have been interesting to look at, such as birth weight and final gestational age (76). Collection of this data should be considered for future studies.

5.4.3. Maternal blood biomarkers and infant DNA methylation

Association between maternal biomarkers and offspring DNA methylation at MEs was investigated, under the hypothesis that availability of methyl-groups could modulate DNA methylation. DNA methylation is important for gene expression regulation and DNA integrity and stability. The association between one-carbon metabolism and DNA methylation has been shown through epidemiological, clinical and animal studies looking at methyl-donors and cofactors availability and disease, but the mechanistic links remain unclear. SAM channels the methyl-group transfer, and deficiency or excess of methyl-donors and co-factors can impair the methyl-group supply and affect the ability to maintain DNA methylation patterns in replicating cells.. The complexity of this process has

already been highlighted in the introduction (Section2.2.2). The existing redundancy of pathways to ensure methyl-group supply (164) for methylation of DNA, amongst other reactions, and the tight regulation of one-carbon metabolism (199) suggest that this is a crucial process requiring a robust system. Yet, the effects of methyl-groups deficiency on DNA methylation have shown to be highly complex: they appear to be dependent on cell type and organ, the stage of cell transformation (e.g. in cancers,) and are gene- and site-specific (200). Most of the studies (including supplementation trials) have investigated in the context of cancer research, where the divergence in direction (i.e. global hypomethylation coexistent with targeted hypermethylation) clearly suggests that this aberrant methylation is not random, although much remains unknown about the regulation. Different studies have been conducted on the effect of depletion/repletion of folate in DNA methylation, but the findings suggest that responses can vary largely depending on age, genotypes, duration and magnitude of the exposure (199).

The same theoretical dependence of DNA-methylation on methyl groups and one-carbon metabolism is expected to happen in the developing fetus, and as a result of the erasure and resetting of the DNA methylation patterns and the rapid growth rate in this period, need for methyl-donors intake could be higher. From a biological point of view, however, it would seem unreasonable that gene methylation changes are so vulnerable to discrete dietary variation, particularly if the genes are involved in important phenotypes in terms of survival. Animal models, nonetheless, seem to indicate the contrary (68, 201, 202). For example, it has been suggested in mice that maternal dietary supplementation may positively affect health and longevity of the offspring (68, 69). Both this potential protection of relevant genes and the success of MEs in animal studies enforce the case for the use of MEs.

Due to the tissue-specificity of one-carbon metabolism, it could be argued that maternal blood concentrations might not be an accurate reflection of the overall maternal

one-carbon metabolism activity or at other maternal tissues, such as the liver (111). However, for this type of epidemiological model, where the focus is in maternal nutritional exposure, blood biomarkers are relatively easy to investigate, compared to other potentially more invasive sampling methods. In addition, the maternal-foetus exchange happens through the placenta via maternal blood, and therefore it would expected that maternal blood concentrations are a good reflection of the maternal availability, and particularly of the potential which might be transferred to the foetus. In fact, it has been shown that maternal SAM, B12 and homocysteine concentrations were associated respectively with infant DNA methylation (78, 166) and foetal cord blood (203). In this sense, further study on the transplacental transport of all the different metabolites into the developing fotus and the foetal pools would help in this area of research, although this might be difficult to do in humans. For example, choline is known to be transported across the placenta through choline transporters (169, 204-206), but betaine and DMG transfer remains unclear. Still, because in this study a sample was obtained only at one time point, the data could be underestimating other existing maternal pools available for use. Nevertheless, it is a useful epidemiological model with relatively easy available maternal biomarker measurements, when measurements in the foetus would be unfeasible to use.

The key findings of the main study are presented in Section 5.1. Most of the biomarkers under study had acceptable levels in this population, with the exception of B2. The severely deficient B2 maternal status observed was associated with infant DNA methylation, when assessing all MEs combined, as were the SAM:SAH and the DMG:betaine ratios, and cysteine. It is important to note that both SAH and DMG are the most challenging biomarkers, due to quite poor reliability and stability and subsequent issues regarding the back-extrapolation to time of conception. Surprisingly, these two biomarkers seem to be the key biomarkers presenting a clearly significant association with DNA methylation, suggesting that this association could be even stronger.

Associations with other biomarkers of known importance, such as folate or choline were not identified in these data. Nevertheless, the data of this study seem to suggest that B2 could be the key limiting factor in this population, and that the more severe the B2 depletion is (higher EGRAC values), the more impaired the DNA methylation is. Additionally, B2 deficiency would impair the folate-dependent way, leading to a switch to the betaine-dependent pathway. The use of this betaine-pathway is also negatively associated (i.e. the higher DMG:betaine ratio) with the DNA methylation, suggesting that this pathway might not be coping with the maintenance of DNA methylation, leading to a lower rate of remethylation of homocysteine, that would go through the transsulfuration pathway to avoid homocysteine accumulation. This could explain also the apparent negative association of cysteine (which is product of the transsulfuration and would increase with a greater use of the transsulfuration pathway) with DNA methylation.

Since the primary assumption was that MEs would respond similarly to methyl-group availability, all MEs were analysed combined. In addition, a second step was taken in the analysis, looking at the associations of biomarkers with individual MEs to identify if they responded differently to biomarkers. This analysis showed considerable differences in the association between biomarkers and the different MEs, with *RBM46* identified as the locus most sensitive to differences in biomarker concentrations. This would suggest that the different loci are not similarly affected by nutritional status and that, therefore, other regulatory / adaptive mechanisms may trigger the response (or lack of response). In fact, during the Dutch Hunger winter study, increases and decreases of DNA methylation as a response to undernutrition were reported, suggesting this possibility (75). When interactions were checked between methylation at the different loci and each of the different biomarkers, and differences tested for each loci, the results were non-significant, indicating that such differences in response of different loci to biomarkers do not exist. It is thus possible that these individual associations shown in Table 5.15 could be just due to

chance (since the associations presented although significant are generally weak). Furthermore, of all of the individual biomarkers, the one which was most frequently associated with DNA methylation was cysteine, with significant associations and the same directionality of association at three loci: *BOLA3*, *RBM46* and *ZNF678*. However, given that cysteine is not on the remethylation pathway but in the transsulfuration pathway, this association could be due to confounding rather than via genuine causation. It is also interesting to note that the *RBM46* sequence analysed contains more than twice the number of CpGs than the other loci sequences. This could lead to higher potential variation and partly explain differences in response.

Finally, the lack of interactions (by pairs of biomarkers) could be due to a real lack of effect, or, if an effect of biomarkers existed, because of a higher complexity of interactions between pathways, as a result of the cyclic and complementary nature of one-carbon metabolism (see Figure 4.1) or due to too small a sample to detect such effects. Indeed, some of the findings regarding associations between biomarkers are inconsistent and appear paradoxical and this needs further research.

Observational studies can be used to detect rapid large effects but also subtle effects so long as there is a suitable source of variation in exposure (106, 107). In this study, changes during development in infant DNA methylation at MEs are thought to happen within a short time frame around conception (79). Also the ranges of exposure observed (i.e folate plasma) are not large but are comparable to other studies (111, 166). Finally, even if the size effect in DNA methylation change necessary for phenotypic alteration remains largely unknown (76), the DNA methylation changes induced by maternal malnutrition have been proved to be detectable (74, 75). Therefore, an observational design such as this one is value, but it is prone to bias and confounding that needs to be considered.

As discussed is Section 3.2.2., the dry season conception group were additionally enrolled within the ENID trial (see Section 5.2.1). It is thus important to consider any possible bias arising from this difference between the groups. Women enrolled into ENID were randomised to either of both intervention groups from booking (<20 weeks gestation) until delivery: either placebo (Iron-folate (FeFol = standard care, given to all participants), energy-protein supplement, multiple micronutrients (including, B2, B6 and B12), or both energy-protein and micronutrients (207). Then their infants were further randomised at six-months of age to a lipid-based nutritional supplement, with or without additional micronutrients. However, for the hypotheses under investigation and the results presented in this thesis (DNA methylation at MEs) supplementation group information was purposely not included into the analysis. DNA methylation at MEs is by definition established around conception and would not be affected by supplementation after 12 weeks. Further, and in relation to the biomarkers, the sample collection in mothers took place prior to any supplementation and thus, maternal biomarkers were not affected by supplementation status. In infants, the blood sample was also collected before postnatal supplementation. However, their DNA methylation patterns outside MEs could be affected by maternal supplementation during pregnancy. Thus, this needs to be taken into account in future analyses (e.g. global methylation, as discussed later in Section 6.3.) though an analysis with four supplementation arms may decrease substantially the power of this study.

In addition to the biomarkers under study, other nutritional/dietary factors have been postulated to affect DNA methylation. Examples are the effects of nutrients through indirect modifications of one-carbon metabolism (e.g. zinc or selenium) and via activity of DNMT enzymes (e.g. polyphenols), and have been reviewed by McKay & Mathers (208). Zinc in The Gambia is not found in high concentrations in foods which are seasonally available (209) but there is no available data showing seasonality in the status of selenium

or polyphenols. Non-nutritional factors, that may also exhibit seasonal confounding could include exposure to aflatoxin or cooking smoke, both of which are known to have the potential to influence DNA-methylation (184, 185, 210). Aflatoxin is a mycotoxin produced by *Aspergillus spp*. and found as contaminant in staple foods in The Gambia and other parts of sub-Saharan Africa, where staples are stored in hot and humid conditions (211). Aflatoxin-DNA adducts have been significantly correlated in Taiwanese hepatocellular carcinoma patients with the DNA methylation of the tumour suppressor gene *RASSF1A* (212), and of *MGMT*, a gene that encodes O^6-methylguanine-DNA methyltransferase (210, 213). If aflatoxin exposure does result in changes in DNA methylation, independent of methyl-group availability, it would be a potential confounder, since the exposure to aflatoxin could also differ considerably between seasons (e.g. recently harvested groundnuts *vs.* long-time stored or imported groundnuts: as the local food availability varies, the sources of food vary too). Regarding the cooking smoke exposure, it was observed within the non-pregnant indicator group study (Section 4.2), that the majority of kitchens were mud brick huts with variable but generally poor ventilation and that wood was used for fuel in 100% of cases. Weather conditions / season did not seem to effect location of cooking (inside / ouside), and thus this exposure seems to be mostly homogeneous throughout the year. A further potential seasonal difference is the workload of the women, which is much more intense during the rainy season, and which might have had an impact on the daily dynamics and indirectly on the total cooking smoke exposure. For example, the time spent cooking, the number of times that they cook a day or even the distribution of the workload at the household amongst the different members of the family (e.g. younger girls and older women) could have the potential to vary considerably.

5.4.4. Conclusions

Evidence for the causal effects of methyl-donors and cofactors on DNA methylation in humans comes mostly from intervention studies on cancer and aging (214, 215) and, overall, DNA methylation at specific loci correlates positively with methyl-group availability (208). However, the duration and the intensity of the exposure necessary to provoke epigenetic changes remains largely unknown, particularly in foetal development, where studies are also required to identify of optimal timings in pregnancy where DNA methylation is sensitive to nutritional exposures. This observational study is subject to confounding and causality cannot be inferred from the data. However, it highlights some important and interesting associations that should be investigated further. Certainly, a more comprehensive investigation should be conducted to assess the association between season and one-carbon metabolites and new DNA methylation outcomes, and provide deeper insights into the determinants of prenatal DNA methylation patterns. In summary, this study presents evidence for associations of infant DNA methylation at MEs with environmental influences (i.e. season), including nutritional maternal condition as assessed by maternal BMI in early pregnancy and SAM:SAH or DMG:betaine ratios, B2 and cysteine. The results from this study are not definitive and conclusive, given its observational nature but clearly support the hypothesis that environmental factors during pregnancy influence the DNA methylation in infants. As such, replication and expansion in this area is vital.

Part III:
GENERAL DISCUSSION

CHAPTER 6: GENERAL DISCUSSION

'We believe fervently that improving nutrition for pregnant women and children under two is one of the smartest investments we or anyone can make' (Hillary Clinton, September 20, 2011)

This Chapter summarises the main findings from this PhD thesis and interprets the findings beyond the scope already covered in earlier sections. The importance of these data within the wider context of health across the life-course is discussed and future research areas arising from this project are suggested.

6.1.SUMMARY

The DOHaD research field has steadily attracted attention over the past years, largely due to the increasing prevalence and burden of NCDs and the global attempts at disease prevention at every life-stage. To develop appropriate prenatal strategies, a better understanding of the biological mechanisms of developmental programming is needed. Among the underlying mechanisms so far proposed, alterations to DNA methylation patterns appears as a strong candidate. Evidence is growing on the influence of nutrition on inter-individual variation in DNA methylation patterns throughout life in humans (208), but the establishment *in utero* of such patterns is still poorly understood. The purpose of the research presented in this thesis was to contribute to the existing evidence base on environmentally-driven epigenetic changes during the human prenatal period.

The research hypothesis under investigation was that season of conception affects DNA methylation and that this can be explained, at least partly, by changes in maternal nutrition-related one-carbon metabolism. The study was designed as two complementary research projects in rural Gambia (the 'indicator' and 'main group' studies), to appraise in combination the relationships between maternal dietary intakes of methyl-donors and cofactors, maternal nutritional status as assessed by biomarkers of one-carbon metabolism and infant DNA methylation at MEs. The blood biomarker data was the common feature linking both studies. The two studies were purely observational and, therefore, not intended to determine causal associations, but should rather be seen as opportunity for hypothesis generation in the enquiry of links between maternal one-carbon status and infant DNA methylation. This is a fast moving field and since this study was designed, further progress has been made in epigenetics (216), such as the analysis of the DNA methylome of different normal (217) or cancerous (218) human cell types. Knowledge on the role of maternal methyl-donors and cofactors in the establishment of offspring DNA methylation patterns has also expanded over the last 3 years (78, 166, 219).

Yet, our understanding remains very limited and much is still to be learned about the complexity of epigenetic regulation (including DNA methylation), thus highlighting the need for further, focused research in this area.

6.1.1. Synopsis of findings

A main finding of this study was that the diet of women of child-bearing age in West Kiang (as assessed in the 'indicator' group) was moderate to severely deficient for several of the methyl-donors and cofactors, namely folate, B2, choline and B6, in comparison to the EARs, whilst intake was adequate for B12 and methionine. Furthermore, intakes of folate, B2, choline and betaine were subject to significant fluctuation throughout the year, whereas methionine, B6 and B12 intakes were stable throughout the year. Patterns of seasonality varied in phase (time of the variation) and amplitude (magnitude of the variation) amongst nutrients under study, the rainy season being richer in some (betaine and B12), and the dry season in others (folate, B2 and B6).

Maternal blood biomarkers were assessed in both sub-studies, as an outcome of season and dietary intake in the indicator group, and as an intermediate exposure in the main group. From the indicator group data, it could be concluded that diet does not translate strongly into blood biomarkers. In addition, reliability and stability were generally good for blood biomarkers, whilst for dietary intakes reliability and stability were limited. Despite the observed deficiencies in dietary intakes, concentrations of blood biomarkers of one-carbon metabolism in Gambian women were mostly adequate. However, important deficiencies in B2 and choline were observed, both in pregnant and non-pregnant women. All blood biomarkers were subject to significant seasonal fluctuation throughout the year, and such seasonal patterns were similar in pregnant and non-pregnant women for several of the biomarkers (folate, choline, betaine, B12, active B12 and DMG:betaine ratio), but differed for the other substances under study. The trends

between seasons for the different biomarkers were similar in non-pregnant and pregnant women, except for choline, methionine, SAM and cysteine. However, absolute biomarker levels were generally lower (except for folate, B2 and SAH) in pregnant compared to non-pregnant women. Of all the biomarkers, special attention was given to the SAM:SAH ratio as an index of methylation. This was based on the hypothesis that the SAM:SAH ratio would be a meaningful summary measure of methyl-groups availability and reflecting maternal methylation capacity, which could be 'transmitted' to the foetus and thus correlating with infant DNA methylation levels. The SAM:SAH ratio was higher during the rainy season in both sub-studies (pregnant and non-pregnant women), suggesting a higher methylation potential at this time of the year.

Finally, previously published findings of an effect of season of conception on infant DNA methylation (79) were successfully replicated in a bigger, independent sample, and expanded with the presentation of three new MEs. A consistent effect of season on the DNA methylation at MEs was observed, with higher methylation levels seen in infants conceived during the rainy season, with the effect size (methylation proportion) and significance varying by locus. Further investigation on likely intermediate exposures highlighted that maternal nutritional status, as assessed by BMI and one-carbon metabolism biomarkers in blood, appeared to affect DNA methylation patterns. Several of the individual biomarkers, namely SAM:SAH ratio, DMG:betaine ratio, B2 and cysteine, as well as all biomarkers jointly (condensed as the first principal component) were associated with infant DNA methylation across all MEs. Although the approach taken does not allow us to establish causality in the relationship between maternal biomarker status and infant DNA methylation, the findings strongly support the hypothesis that maternal nutritional status plays an important role in early epigenetic processes in the offspring.

6.2.DISCUSSION

The final aim of this thesis was to identify whether links exist between the periconceptional seasonal maternal diet and infant DNA methylation (objective 5, Section 2.5.). For ethical and practical reasons (as discussed in Section 5.3.1) the dietary intake was only assessed in the indicator group of non-pregnant women and could thus not be directly related to infant DNA methylation. The measurements were however thought to represent a good estimation of dietary intake of women in rural Gambian including, importantly, at time of conception. As outlined earlier, intakes of betaine and B12 were higher during rainy season, and diet during the dry season was richer in folate, B2 and B6. Dietary intake levels do mostly not agree well with the findings on blood levels of these substances in non-pregnant or pregnant women. Tables 4.2 and 5.3 show that blood betaine, folate and B2 were higher during the rainy season, and conversely, B6 and B12 were higher during the dry season. Correlations between dietary intake and blood biomarker levels could be tested formally in the indicator group. This confirmed that betaine, B6, B2 and B12 intakes were associated with the levels in blood, although the effect size (regression coefficients) was not as strong as expected. This lack of consistent correlations and the limited reliability and stability of the dietary intake data suggests that the biomarker data is preferable over the dietary intake data for this type of study, at least in this population.

Regarding biomarkers, the high interdependence of one-carbon metabolites means that different 'road blocks' can impact on the biomarker profile and lead to different scenarios according to limiting metabolites. For example, severe deficiency of B2 in this population might impair the biological availability or metabolic utilization of another nutrient (e.g. within the folate pathway), but this limitation might be buffered by the availability of yet another metabolite (e.g. those in the betaine-pathway) (220). This also

suggests that associations between biomarkers could be different in other populations with different dietary deficiencies and intake patterns.

Furthermore, in view of the intra-individual biological variability in the blood biomarkers (118), which could have been further exacerbated by sample handling (particularly processing time due to the field conditions), the risk of using one single time point of measurement should also be considered. The plasma concentrations measured give information of a specific moment (and thus reflecting only partial view of a complex process), without information on flux or turnover and ignoring conversion rates and allosteric regulation. Most of the metabolites involved in one-carbon metabolism make up part of a bigger body pool (e.g. body proteins) and also play other roles in the body (e.g. choline function in cell membrane or brain development). This makes the blood levels perhaps unrepresentative of the real status. Analysis of serial samples or studies looking at metabolite kinetics could be more illustrative of the actual biomarker status and methylation potential over a period of time (i.e. in early pregnancy). In addition, less invasive methods would be more appropriate for this kind of sampling, particularly in participants who have not yet been confirmed as pregnant. The availability of appropriate laboratory techniques based, for example on urine would be an important improvement in this respect.

This study included the assessment of a comprehensive set of biomarkers and focused on methyl-donors and cofactors, as well as the functional biomarkers of one-carbon metabolism. However, the list of substances of interest could be expanded further for a better understanding within the context of the overall demands for one-carbon units. For example, a substantial competitive demand for methyl-groups exists from endogenous synthesis of substances such as creatine (221). In case of shortage in the availability of methyl-groups, constrictions on the synthesis of such substances might spare methyl-groups for DNA methylation. Creatine acts as a storage fuel for skeletal muscle and might

be associated with body weight and lean mass and thus with seasonality (222). In this setting, patterns in physical activity are very dependent on the season, particularly with women working intensively the fields during the rainy season (153), and this could potentially affect the muscle mass. Furthermore, other nutrients such as serine and glycine are involved in the methyl supply to the folate-dependent pathway (see Figure 4.1), which were not measured, as well as measures of the transsulfuration cycle (e.g. cysteine glutathione).

The nature of one-carbon metabolism is an unusually complex network, with redundancy of interlocking cyclic pathways and ongoing recycling of metabolites (e.g. the remethylation of homocysteine, the endogenous synthesis of choline (111, 223)). Understanding the highly intricate nutrient-nutrient interactions within this context requires a comprehensive approach and, ideally, a larger sample size. Deeper analysis of higher level interactions (i.e. involving several of the biomarkers) could be done using a sophisticated modelling approach which was beyond the scope of the current thesis, but may be explored in future analyses.

The inter-individual variations in DNA methylation epigenotype at birth can be explained by environmental, genetic or stochastic factors (78, 111). However, at MEs, by definition, epigenotype is not genetically determined, but established stochastically and is sensitive to environment (224). Infant DNA methylation was generally higher in the group of infants conceived during the rainy season, as seen in an earlier study within the same population setting (79). However, compared to previously reported findings the methylation proportions and the statistical significance of observed differences between seasons were weaker. Furthermore DNA methylation levels, irrespective of season, were observed to be lower in the present study. The reasons why this could be the case are discussed in detail in Chapter 5, with the most plausible explanation being that secular trends between the time points that the infants under investigation were conceived

(1991-1998 and 2009-10) has diminished seasonal differences. Yet, seasonal differences in methylation patterns still exist, and thus this present work represents a replication, albeit at reduced statistical significance, in this rural Gambian population. The association of DNA methylation at MEs with the principal component of one-carbon metabolism biomarkers, which provides with a helpful means for the reduction of random noise in complex data, can be difficult to interpret (225). Nonetheless, DNA methylation was also associated with availability of individual nutrients such as B2. This would seem logical, given that B2 is the most limiting nutrient under study, although in theory this deficiency could potentially be compensated by the betaine-dependent pathway. DNA methylation was also associated with plasma cysteine and with the two ratio measures, SAM:SAH and DMG:betaine, employed as indicators of the transmethylation (or transfer of methyl-groups for biological methylation) and betaine-dependent remethylation pathways, respectively. However, these biomarkers and ratios are all correlated with season, so it is difficult to ascertain, for example, whether the SAM:SAH ratio explains the seasonal effect on DNA methylation or whether season explains the relationship of DNA methylation with SAM:SAH (i.e. the association of DNA methylation with SAM:SAH ratio might be spurious and simply due to its correlation with season or with some other seasonal unmeasured factor exerting the true effect on DNA methylation). Despite the various considerations in the interpretation of the data outlined above, this study provides us with some initial evidence of associations of infant DNA methylation at MEs with prenatal environment and maternal nutritional status, which precipitates the need for further study. Other studies have found associations of DNA methylation with different prenatal one-carbon metabolism biomarkers (78, 166). For example, Jing et al. (219) have recently reported that higher maternal choline intakes (930 *vs.* 480 mg/d) in the third trimester of pregnancy may modify foetal epigenetic state of cortisol-regulating genes and their expression in the placenta (219). In this study, choline affected placental gene promoter

methylation and was associated with higher global DNA methylation in the placenta (219). There is also evidence of genetic components to the regulation of DNA methylation patterns (208). In addition, other factors could also be involved in the effect of season of conception on DNA methylation (i.e. factors explaining adaptive responses with broader physiological response to a difficult environment, where low methyl-groups availability could perceived as limiting).

The project described in this thesis has been very successful at: i) surveying and sampling every month, for one calendar year, 30 women with a high degree of detail (indicator group), and ii) screening and following up the menstrual cycle of 1630 women to identify eligible women, conceiving within the established timeframe. This is quite remarkable in the given setting, and must be viewed as a strength of the study, which required an enormous logistic effort. This was made possible by coordinating resources with other projects and utlizing the existing research platforms (e.g. the DSS) at MRC Keneba. With respect to the findings described, this study has produced novel data on a comprehensive set of one-carbon metabolites and co-factors in terms of both dietary intakes and plasma biomarker levels. Such data were not available for this population before. The study also provided information on Gambian indigenous food concentrations of substances under study, which contributes to the relatively limited food composition information existing on West African foods and African foods more generally and, in the case of choline and betaine, to the global database, recently started by USDA (141).

MEs are thought to be established very early in gestation (224). Therefore the ideal time to measure nutritional exposure would be around conception. The study samples were collected up to 16 weeks post-conception. Hence, it was deemed imperative to correct the biomarker concentrations in the main group women for seasonality, after assessing whether the indicator group is representative of the main group, and the measurement error (stability and reliability). The limitation of this back-extrapolation

process has been discussed in Chapters 4 and 5. Given the environment in which the research was conducted, collecting blood samples closer to conception will always be a challenge, unless different sampling methods are used, as discussed above, based perhaps on saliva or urine, with less processing required, and which would allow for blanket sample collection (less resource heavy than venous sampling), as well as being more discrete and allowing more frequent sampling. Such alternative methods need further exploration. However, such developments, and the interpretation of data obtained, would need to consider the issue of tissue specificity of one-carbon metabolism. Blood concentration of biomarkers might be a limited reflection of the general maternal methylation potential in the different tissues but it is more likely to reflect what biomarker levels are available to the foetus via placental transfer.

In addition, the 16 weeks cut-off for early pregnancy blood sample collection led to some participants conceiving within the *a priori* specified time frame to be excluded from the study. This reduced the sample size and thus the power for detection of association between the exposures and outcomes, especially in view of multiple testing. The risk exists that associations identified, and these often not being highly significant, could be due to chance. Clear statements about allowances for this were made, e.g. the application of the Bonferroni correction to some of the analyses. There is also potential for an inability to detect true differences, and thus committing type II errors, particularly in the investigation of the association between individual biomarkers and DNA methylation, and the interactions between biomarkers.

6.3.FUTURE RESEARCH

This thesis is based upon the research conducted within the Gambian MDEG project, one of the first studies to look at effects of maternal one-carbon metabolism in the establishment of DNA methylation patterns during early development, in a comprehensive manner. Using the MDEG study, we continue to ask further research questions, as described in the following section. The initial results obtained so far are of limited applicative use, but raise ideas for further studies in this area as outlined below.

6.3.1. Ongoing research within the MDEG framework

Within the framework of the existing study, there are several immediate next steps planned and underway using available samples, with the aim to enhance our understanding of the data obtained so far. For example, genetic variation in genes involved in one-carbon metabolism could add an extra layer of complexity to the interpretation of the results. Genetic variation (e.g. single nucleotide polymorphisms or SNPs) can substantially affect one-carbon metabolism, leading to the modification of dietary requirements. For example, SNPs in the MTHFR gene can lead to a 30-60% reduced efficiency of the enzyme methylenetetrahydrofolate reductase, which affects folate and homocysteine concentrations (226). Similarly, a polymorphism (rs12325817) in the PEMT gene (phosphatidylethanolamine-N-methyltransferase), involved in endogenous liver choline production and induced by oestrogen, exists. Almost 75% of the population of North Carolina in the US has one variant allele for this SNP, which increases dietary requirements for choline (136, 227). However, the frequency of such variants in African populations remains mostly unknown. In collaboration with Prof. Steve Zeisel (University of North Carolina, USA), we are analysing the maternal DNA to establish the frequency distribution of a panel of approximately 300 SNPs in about 18 genes involved in one-carbon metabolism. This information should expand our understanding of inter-individual

variation in levels of one-carbon metabolites and potentially help interpret our data even with respect to infant DNA methylation patterns.

Investigating cross-cultural data in maternal one-carbon metabolism may help understand how it adjusts to specific deficiencies within individual population groups. Contemporarily to the analyses that were conducted in the MDEG maternal plasma samples, identical analyses were carried out by our collaborator, Prof Sheila Innis (University of British Colombia, Canada), on a cohort of women of reproductive age from Canada. A comparative analysis of biomarker concentrations will be undertaken between women from The Gambia (both pregnant and non-pregnant) and Canada (non-pregnant). In Canada mandatory folate fortification has been implemented since 1998 (with fortification of 150 μg of folic acid per 100 grams of enriched flour and uncooked cereal grains). The Canadian dataset also includes information on dietary intake of the methyl-donors and cofactors, which thus can be compared to data from the indicator group of women in The Gambia. Since fortification was introduced in Canada, a dramatic increase in blood levels of folate and a substantial decrease in plasma homocysteine levels has been observed in the population (228). Our unsupplemented population is likely to respond more dramatically to fluctuations in intakes, particularly in pregnancy. Additionally, a third cohort of pregnant women from The Netherlands could be investigated, where similar biomarker analyses were carried out (175). The Netherlands represents another high-income country, but without mandatory folate fortification. A three-way comparison of data between Gambian, Canadian and Dutch data could thus shed light on the role of dietary intakes (in this case deficiencies and supplementation) of one-carbon metabolites and cofactors, as well as pregnancy stage, on blood biomarker levels.

In terms of DNA methylation, MEs are an obvious target to investigate DNA methylation patterns due to their particular characteristics of environmental sensitiveness, and establishment in early development that remains thereafter, across

tissues. However, a whole genome approach can be considered both complementary and more comprehensive than a locus-specific approach. Moreover, the establishment of methylation patterns genome-wide may respond differently to methyl-group availability than MEs and be subject to different timing (de- and re-methylation may occur over time at loci not considered to be MEs). Potentially, genome-wide methylation is less tightly regulated than methylation of MEs and thus more sensitive to maternal variation in one-carbon metabolism. It is not unusual, for example in cancers, to observe global hypomethylation coexisting with hypermethylation at specific sites (229). Therefore a genome-wide analysis could lead to interesting results complementary to the ones presented, particularly if the analyses covered the same loci or CpGs described as part of this thesis. The analysis of global DNA methylation for the infant DNA samples is currently ongoing (using the Illumina 450K Infinium BeadChip) at the International Agency for Research on Cancer (IARC) and in collaboration with colleagues at the University of Leeds. The primary aim of this research is to investigate whether maternal aflatoxin exposure (plasma aflatoxin-albumin adducts assessed in MDEG maternal plasma samples) affects infant DNA methylation. Since the exposure to aflatoxins has previously been reported to be seasonal in The Gambia (230, 231), and if it shows to have an effect on infant DNA methylation of this cohort, the aflatoxin exposure data could be used to control the current analyses, as it could an important source of confounding between season and infant DNA methylation.

To progress further the research of MEs with the Gambian setting, work will continue in Rob Waterland's lab in order to: identify new MEs; verify already identified MEs in different tissues in Gambians; evaluate potential phenotypic effects of MEs in the infants in follow-up studies.

Finally and most importantly, the ultimate interest of an effect of maternal diet on infant DNA methylation is based on the assumption that this can affect the health

outcomes of the offspring long-term. Further research is necessary to establish if there are any biological consequences of altered DNA methylation patterns. Identifying differential methylation at genes associated with specific phenotypes such as growth or early markers of NCDs could help greatly to elucidate potential mechanisms underlying the DOHaD hypothesis. Since a strong asset of this research setting is the ability to track and follow-up study participants, a new add-on study to MDEG has been granted ethical approval, for which we have initiated the fieldwork. We are now collecting blood samples of the study children when they reach 2 years of age, together with detailed anthropometric data. The specific panel of phenotypic outcomes to be assessed in the blood samples collected, still needs to be defined in detail. As an example, the PAX8 gene, one of the MEs investigated, is involved in hypothyroidism and future work could investigate as to whether methylation at this ME has phenotypical consequences on thyroid function (e.g. levels of thyroid stimulating hormone and free thyroid hormones FT3 and FT4). Similarly, the hypothalamic-pituitary-adrenal axis that regulates cortisol production, has been shown to be sensitive to perinatal epigenetic programming (219) and could therefore be a good system to investigate, in order to assess potential consequences of alterations in DNA methylation patterns. Data suggests that measures of immunity may be modulated by season in this environment (99) and thus immunological outcomes can also be considered in the context of epigenetic analyses. Furthermore, phenotypes related to growth and IGF-axis, and/or early markers of diabetes would also be interesting to investigate (however diabetes prevalence is very low in rural Gambia). Remaining DNA stocks could be used as basis for in-depth methylation assessment at relevant key genes, e.g. IGF genes, which are involved in regulating immune function and growth, if deemed appropriate.

6.3.2. Potential future research within the MDEG framework

As indicated above several lines of research are already under way expanding on the work described in this thesis. Still, much additional work could be conducted within the framework of the existing study. For example, another outcome to consider is DNA methylation at imprinted genes. Imprinted genes control fetal and placental growth (among other things) and are involved in the adaptative response to the intra-uterine environment (232). Their expression is exclusively from one parental allele. This monoallelic expression is initiated by differential methylation between the oocyte and sperm in differentially methylated regions (DMR). These DMRs are thought to be particularly susceptible to nutritional deficiencies around conception (105). Periconceptional micronutrient supplementation seemed to alter offspring DNA methylation in a Gambian population (105), although these results were based on a small sample, and these genes could be investigated in much greater detail in this present cohort.

As already discussed, one-carbon metabolism is complex and the interpretation of the results obtained is limited by the static nature of the data obtained as a part of this thesis. Sophisticated mathematical modelling, with data on related enzyme kinetics and regulatory mechanisms could help our understanding of the data available. Scientists at the Duke University, Durham, US, have developed mathematical simulations of one-carbon metabolism with predictions that match experimental data (223) and an in silico experimentation model has been designed from those studies (233). We are exploring a collaboration with this group to run our data initially through their existing in silico model, albeit with the caveat that models developed on Caucasians might not be the best fit for our Gambian population

Another area worth to investigate would be the evolution in the maternal plasma concentrations of the biomarkers of interest throughout pregnancy. Although only one

maternal sample was collected for MDEG study women, the ENID Trial collects samples at booking (< 20 weeks), 20 and 30 weeks gestation. Plasma samples from women who were on placebo could thus be analysed at the three consecutive time points to assess the intra-individual variation throughout pregnancy, which would coexist with the seasonal variation.

6.3.3. Further research on one-carbon metabolism and DNA methylation *in utero* beyond the MDEG framework

Although the association of season of conception with altered DNA methylation was replicated in this current study, further corroboration of this type of observational study in different human populations is optimal, particularly in less developed countries with marked nutritional limitations (e.g. Indian populations, with good folate status and yet a noted B12 deficiency (138)).

Regarding the possible role of one-carbon metabolites in DNA methylation in utero, the strongest evidence available for a direct effect so far comes from supplementation studies of methyl-donors in murine models and shows both upregulation (234) and downregulation of DNA methylation (235, 236). To explore in further detail the physiology of one-carbon metabolism related to our findings (i.e. the effects on the different pathways when availability of specific methyl-donors and cofactors is modified), this research should be complemented with animal work. Animal models would allow a specific focus on responses within certain states of severe deficiency, such as B2 deficiency (which appeared the most limiting factor within our Gambian population). The mouse genomes have a 85% genes of similarity to human genes, making it a powerful and useful model (15). However, there are clearly some challenges when inferring from mice trials in laboratory conditions to observational studies in a free-living human population.

The ultimate goal of this area of research should be, however, to gather enough evidence about the effect of maternal nutrition in one-carbon metabolism in humans to assess optimal maternal needs. Such evidence can only be obtained via randomised controlled trials of nutritional supplementation. Published data from trials conducted in developing countries in pregnant women have included B-vitamins (237) that could be followed up to look at the DNA methylation in the offspring. Research on the DNA methylation of the offspring of supplemented women has already being conducted in the Gambia (104, 105). Follow-up of the participants in these studies to assess phenotypic outcomes of growth and early markers of chronic disease or other phenotypes (as discussed in Section 6.3.1), could be important. So far, supplementation trials in LMICs have not included choline and/or betaine. This study suggests that choline and B2 status in this population are very deficient and that DNA methylation is associated with the betaine-dependent pathway and with B2 (as observed by significance of association with DMG:betaine ratio and B2 ratio). In addition, evidence is growing on the importance of choline, as suggested by the oestrogen-induced production of phosphatidylcholine in the liver (238), the high transplacental transport and high choline levels in breast milk (239). Choline is thought to influence hippocampal neurogenesis during fetal development and thus memory function in infants lifelong (240-244), partly mediated by changes in DNA and histone methylation (241, 245, 246). However, these data need to be supported by firm evidence from trials in humans. The initial results of the MDEG study, complemented by the planned steps, could provide the basis for a potential choline (and betaine) intervention study in The Gambia assessing, for example, infant brain development as an outcome measure. Supplementation with B2 should also be considered. However, caution is needed until we fully understand the impact of supplementation with one one-carbon substances on differing scenarios of status of the other metabolites. We would expect

supplementation studies starting even before conception to achieve the greatest effect on outcome.

6.3.4. Wider context

The establishment of DNA methylation patterns during development is a complex area of research, where human evidence is scarce and challenging to obtain due to the life-course approach needed (from pre-pregnancy to offspring adult life). The results of this study, whilst exploratory, should stimulate future research into this topic. Ideally, as discussed in 6.3.3., randomised controlled trials of nutritional supplementation are needed. Should these prove informative, populations with deficient diets are likely to benefit most from such interventions (and potentially with less risks of side-effects) than those with adequate intakes. Furthermore, particular populations or population subgroups with a high frequency of functional genetic variants affecting the dietary requirements (227, 244) could benefit from targeted interventions. New opportunities for optimising intakes in pregnancy could also be opened up, by translating relevant findings into diet-based interventions.

How *in utero* epigenetic marks are first established and then modulated by subsequent exposures throughout the life-course is also key for the understanding of foetal programming and potentially for therapeutic research (247). The DOHaD hypothesis is now generally accepted (248), and has led to public awareness of the importance of the prenatal period as a critical window in a life-course approach to health, exemplified by the '1000 days initiative' (http://www.thousanddays.org). Decision-makers do now recognise the importance of nutrition during pregnancy and the need for a better understanding, to define the specific relevant pathways and help identify useful pregnancy strategies for optimal child outcomes that can be scaled up and translated into public health policies.

In parallel, the growing evidence of the role of DNA methylation in disease development, and particularly cancer (216), renders this a topic of global importance. This is reflected by enhanced research opportunities in this field. Funding bodies all over the world have been launching epigenetic funding opportunities (e.g. US National Institutes of Health, Canadian Epigenetics, Environment and Health Research Consortium, etc.), and in September 2011,the health-research division of the European Commission launched the BLUEPRINT project, its largest-ever project with a €30-million investment human epigenome research (249). Given that the epigenome seems particularly sensitive during early development, research at this stage is critical and could contribute to effective early disease prevention.

6.4.CONCLUSION

This thesis supports that infant DNA methylation at MEs is sensitive to the effect of season of conception in rural Gambia. It also suggests that this effect is associated with maternal nutritional status of one-carbon metabolites. However, this association cannot easily be attributed to specific nutrients / biomarkers, and further research is required. Given the complexities highlighted, this study reflects a good first attempt at investigating seasonality of methyl-donors in this population and their role in DNA methylation and, as indicated, it provides a departure point for future work. Since this research field is of potential importance for health throughout the life-course, further investigation into the impact of environmental exposures and timing of their maximal effect is necessary, both in The Gambia and elsewhere, for the benefit of future generations.

BIBLIOGRAPHY

1. Hanson M, Gluckman P. Developmental origins of noncommunicable disease: population and public health implications. Am J Clin Nutr. 2011;94(6 Suppl):1754S-8S.
2. Painter RC, Roseboom TJ, Bleker OP. Prenatal exposure to the Dutch famine and disease in later life: an overview. Reprod Toxicol. 2005;20(3):345-52.
3. Lartey A. Maternal and child nutrition in Sub-Saharan Africa: challenges and interventions. Proc Nutr Soc. 2008;67(1):105-8.
4. Wardlaw T, Blanc A, Zupan J, Ahman E. Low birthweight: Country, regional and global estimates. WHO/UNICEF, 2004.
5. Popkin BM, Adair LS, Ng SW. Global nutrition transition and the pandemic of obesity in developing countries. Nutr Rev. 2012;70(1):3-21.
6. Strong K, Mathers C, Leeder S, Beaglehole R. Preventing chronic diseases: how many lives can we save? Lancet. 2005;366(9496):1578-82.
7. Uauy R, Kain J, Corvalan C. How can the Developmental Origins of Health and Disease (DOHaD) hypothesis contribute to improving health in developing countries? Am J Clin Nutr. 2011;94(6 Suppl):1759S-64S.
8. Mukherjee D, Patil CG. Epidemiology and the global burden of stroke. World neurosurgery. 2011;76(6 Suppl):S85-90.
9. Hales CN, Barker DJ. Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. Diabetologia. 1992;35(7):595-601.
10. Gluckman PD, Hanson MA, Bateson P, Beedle AS, Law CM, Bhutta ZA, et al. Towards a new developmental synthesis: adaptive developmental plasticity and human disease. Lancet. 2009;373(9675):1654-7.
11. Gluckman PD, Hanson MA. The developmental origins of the metabolic syndrome. Trends in endocrinology and metabolism: TEM. 2004;15(4):183-7.
12. Langley-Evans SC. Developmental programming of health and disease. Proc Nutr Soc. 2006;65(1):97-105.
13. Tarry-Adkins JL, Ozanne SE. Mechanisms of early life programming: current knowledge and future directions. Am J Clin Nutr. 2011;94(6 Suppl):1765S-71S.
14. Hassler MR, Egger G. Epigenomics of cancer-emerging new concepts. Biochimie. 2012.
15. Microbial Genomics UDoE. Available from: (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml).
16. Finishing the euchromatic sequence of the human genome. Nature. 2004;431(7011):931-45.
17. Holliday R. Epigenetics: a historical overview. Epigenetics. 2006;1(2):76-80.
18. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003;33 Suppl:245-54.

19. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr. 2007;27:363-88.
20. Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KV, et al. Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. Proc Natl Acad Sci U S A. 2003;100(5):2538-43.
21. Morgan DK, Whitelaw E. The role of epigenetics in mediating environmental effects on phenotype. Nestle Nutr Workshop Ser Pediatr Program. 2009;63:109-17; discussion 17-9, 259-68.
22. Guil S, Esteller M. DNA methylomes, histone codes and miRNAs: tying it all together. Int J Biochem Cell Biol. 2009;41(1):87-95.
23. Zeisel SH. Epigenetic mechanisms for nutrition determinants of later health outcomes. Am J Clin Nutr. 2009;89(5):1488S-93S.
24. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 2008;9(6):465-76.
25. Kim JK, Samaranayake M, Pradhan S. Epigenetic mechanisms in mammals. Cell Mol Life Sci. 2009;66(4):596-612.
26. Holliday R. DNA methylation and epigenotypes. Biochemistry (Mosc). 2005;70(5):500-4.
27. Jones PA, Liang G. Rethinking how DNA methylation patterns are maintained. Nat Rev Genet. 2009;10(11):805-11.
28. Bird A. Perceptions of epigenetics. Nature. 2007;447(7143):396-8.
29. Metivier R, Gallais R, Tiffoche C, Le Peron C, Jurkowska RZ, Carmouche RP, et al. Cyclical DNA methylation of a transcriptionally active promoter. Nature. 2008;452(7183):45-50.
30. Laird CD, Pleasant ND, Clark AD, Sneeden JL, Hassan KM, Manley NC, et al. Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. Proc Natl Acad Sci U S A. 2004;101(1):204-9.
31. Clouaire T, Stancheva I. Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin? Cell Mol Life Sci. 2008;65(10):1509-22.
32. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99(3):247-57.
33. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992;69(6):915-26.
34. Peterson CL, Laniel MA. Histones and histone modifications. Curr Biol. 2004;14(14):R546-51.

35. Delcuve GP, Rastegar M, Davie JR. Epigenetic control. J Cell Physiol. 2009;219(2):243-50.
36. Hebbes TR, Thorne AW, Crane-Robinson C. A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J. 1988;7(5):1395-402.
37. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis. 2010;31(1):27-36.
38. Groth A, Rocha W, Verreault A, Almouzni G. Chromatin challenges during DNA replication and repair. Cell. 2007;128(4):721-33.
39. Szymanski M, Barciszewska MZ, Zywicki M, Barciszewski J. Noncoding RNA transcripts. J Appl Genet. 2003;44(1):1-19.
40. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet. 2004;5(7):522-31.
41. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009;19(1):92-105.
42. Cortez CC, Jones PA. Chromatin, cancer and drug therapies. Mutat Res. 2008;647(1-2):44-51.
43. Portoso M, Cavalli G. The role of RNAi and noncoding RNAs in Polycomb mediated control of gene expression and genomic programming. In: Morris K, editor. RNA and the regulation of gene expression: a hidden layer of complexity: Caister Academic Press; 2008. p. 29-44.
44. Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, et al. Dynamic regulation of nucleosome positioning in the human genome. Cell. 2008;132(5):887-98.
45. Burdge GC, Lillycrop KA. Nutrition, epigenetics, and developmental plasticity: implications for understanding human disease. Annu Rev Nutr. 2010;30:315-39.
46. Vaissiere T, Sawan C, Herceg Z. Epigenetic interplay between histone modifications and DNA methylation in gene silencing. Mutat Res. 2008;659(1-2):40-8.
47. Chuang JC, Jones PA. Epigenetics and microRNAs. Pediatr Res. 2007;61(5 Pt 2):24R-9R.
48. Bernstein E, Allis CD. RNA meets chromatin. Genes Dev. 2005;19(14):1635-55.
49. Whitelaw NC, Whitelaw E. How lifetimes shape epigenotype within and across generations. Hum Mol Genet. 2006;15 Spec No 2:R131-7.
50. Sandovici I, Leppert M, Hawk PR, Suarez A, Linares Y, Sapienza C. Familial aggregation of abnormal methylation of parental alleles at the IGF2/H19 and IGF2R differentially methylated regions. Hum Mol Genet. 2003;12(13):1569-78.
51. Moore T, Haig D. Genomic imprinting in mammalian development: a parental tug-of-war. Trends Genet. 1991;7(2):45-9.
52. Sandovici I, Naumova AK, Leppert M, Linares Y, Sapienza C. A longitudinal study of X-inactivation ratio in human females. Hum Genet. 2004;115(5):387-92.

53. Heard E. Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. Curr Opin Genet Dev. 2005;15(5):482-9.
54. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. Hum Mol Genet. 2005;14 Spec No 1:R47-58.
55. Kelly TL, Trasler JM. Reproductive epigenetics. Clin Genet. 2004;65(4):247-60.
56. Junien C, Nathanielsz P. Report on the IASO Stock Conference 2006: early and lifelong environmental epigenomic programming of metabolic syndrome, obesity and type II diabetes. Obes Rev. 2007;8(6):487-502.
57. Hitchins MP, Moore GE. Genomic imprinting in fetal growth and development. Expert Rev Mol Med. 2002;4(11):1-19.
58. Marchal R, Chicheportiche A, Dutrillaux B, Bernardino-Sgherri J. DNA methylation in mouse gametogenesis. Cytogenet Genome Res. 2004;105(2-4):316-24.
59. Chmurzynska A. Fetal programming: link between early nutrition, DNA methylation, and complex diseases. Nutr Rev.68(2):87-98.
60. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A. 2005;102(30):10604-9.
61. Sandovici I, Kassovska-Bratinova S, Loredo-Osti JC, Leppert M, Suarez A, Stewart R, et al. Interindividual variability and parent of origin DNA methylation differences at specific human Alu elements. Hum Mol Genet. 2005;14(15):2135-43.
62. Chen T, Hevi S, Gay F, Tsujimoto N, He T, Zhang B, et al. Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. Nat Genet. 2007;39(3):391-6.
63. Jackson-Grusby L, Beard C, Possemato R, Tudor M, Fambrough D, Csankovszki G, et al. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat Genet. 2001;27(1):31-9.
64. Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, et al. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol. 2008;6(1):e22.
65. Waterland RA. Is epigenetics an important link between early life events and adult disease? Horm Res. 2009;71 Suppl 1:13-6.
66. Herceg Z. Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. Mutagenesis. 2007;22(2):91-103.
67. Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. Nat Genet. 1999;23(3):314-8.
68. 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(11):949-57.

69. Waterland RA, Jirtle RL. Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. Nutrition. 2004;20(1):63-8.
70. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, et al. Epigenetic programming by maternal behavior. Nat Neurosci. 2004;7(8):847-54.
71. Warnecke PM, Stirzaker C, Song J, Grunau C, Melki JR, Clark SJ. Identification and resolution of artifacts in bisulfite sequencing. Methods. 2002;27(2):101-7.
72. Mathers JC. Session 2: Personalised nutrition. Epigenomics: a basis for understanding individual differences? Proc Nutr Soc. 2008;67(4):390-4.
73. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J.
74. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A. 2008;105(44):17046-9.
75. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, et al. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum Mol Genet. 2009;18(21):4046-53.
76. Tobi EW, Heijmans BT, Kremer D, Putter H, Delemarre-van de Waal HA, Finken MJ, et al. DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. Epigenetics. 2011;6(2):171-6.
77. Steegers-Theunissen RP, Wathen NC, Eskes TK, van Raaij-Selten B, Chard T. Maternal and fetal levels of methionine and homocysteine in early human pregnancy. Br J Obstet Gynaecol. 1997;104(1):20-4.
78. McKay JA, Groom A, Potter C, Coneyworth LJ, Ford D, Mathers JC, 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(3):e33290.
79. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. PLoS Genet. 2010;6(12):e1001252.
80. Center UAS. The Gambia Available from: http://www.international.ucla.edu/africa/countries/article.asp?parentid=96744.
81. WorldBank. The Gambia. 2011; Available from: http://data.worldbank.org/country/gambia - cp_wdi.
82. Klugman J. Human Development Report 2010- The real wealth of nations: Pathways to human development (20th anniversary edition). Palgrave Macmillan: UNDP, 2010.
83. UNICEF. The state of the world's children 2011-Adolescence, an age of opportunity. UNICEF, 2011.

84. Siervo M, Grey P, Nyan OA, Prentice AM. Urbanization and obesity in The Gambia: a country in the early stages of the demographic transition. Eur J Clin Nutr. 2006;60(4):455-63.
85. van der Sande MA, Milligan PJ, Nyan OA, Rowley JT, Banya WA, Ceesay SM, et al. Blood pressure patterns and cardiovascular risk factors in rural and urban gambian communities. Journal of human hypertension. 2000;14(8):489-96.
86. van der Sande MA, Ceesay SM, Milligan PJ, Nyan OA, Banya WA, Prentice A, et al. Obesity and undernutrition and cardiovascular risk factors in rural and urban Gambian communities. American journal of public health. 2001;91(10):1641-4.
87. Roberts SB, Paul AA, Cole TJ, Whitehead RG. Seasonal changes in activity, birth weight and lactational performance in rural Gambian women. Trans R Soc Trop Med Hyg. 1982;76(5):668-78.
88. Mc Crae JE, Paul AA. Foods of Rural Gambia. Second ed. Centre MRCDN, editor1996.
89. Poppitt SD, Prentice AM, Jequier E, Schutz Y, Whitehead RG. Evidence of energy sparing in Gambian women during pregnancy: a longitudinal study using whole-body calorimetry. Am J Clin Nutr. 1993;57(3):353-64.
90. Gilham S. An Investigation of factors influencing dietary intake in West Kiang, the Gambia. MSc Summer project: LHSMT, 2009.
91. Prentice A, Laskey MA, Shaw J, Hudson GJ, Day KC, Jarjou LM, et al. The calcium and phosphorus intakes of rural Gambian women during pregnancy and lactation. Br J Nutr. 1993;69(3):885-96.
92. Villard L, Bates CJ. Dietary intake of vitamin A precursors by rural Gambian pregnant and lactating women. Human nutrition Applied nutrition. 1987;41(2):135-45.
93. Prentice AM. The emerging epidemic of obesity in developing countries. Int J Epidemiol. 2006;35(1):93-9.
94. Whitehead RG, Rowland MG, Hutton M, Prentice AM, Muller E, Paul A. Factors influencing lactation performance in rural Gambian mothers. Lancet. 1978;2(8082):178-81.
95. Bates CJ, Prentice AM, Paul AA. Seasonal variations in vitamins A, C, riboflavin and folate intakes and status of pregnant and lactating women in a rural Gambian community: some possible implications. Eur J Clin Nutr. 1994;48(9):660-8.
96. Prynne CJ, Paul AA. Food Composition Table for use in The Gambia. MRC Human Nutrition Research, 2011.
97. Moore SE, Mansoor MA, Bates CJ, Prentice AM. Plasma homocysteine, folate and vitamin B(12) compared between rural Gambian and UK adults. Br J Nutr. 2006;96(3):508-15.
98. Brewster DR, Greenwood BM. Seasonal variation of paediatric diseases in The Gambia, west Africa. Annals of tropical paediatrics. 1993;13(2):133-46.

99. Moore SE, Collinson AC, Tamba N'Gom P, Aspinall R, Prentice AM. Early immunological development and mortality from infectious disease in later life. Proc Nutr Soc. 2006;65(3):311-8.
100. Moore SE. Nutrition, immunity and the fetal and infant origins of disease hypothesis in developing countries. Proc Nutr Soc. 1998;57(2):241-7.
101. Rayco-Solon P, Fulford AJ, Prentice AM. Differential effects of seasonality on preterm birth and intrauterine growth restriction in rural Africans. Am J Clin Nutr. 2005;81(1):134-9.
102. Prentice AM, Cole TJ, Foord FA, Lamb WH, Whitehead RG. Increased birthweight after prenatal dietary supplementation of rural African women. Am J Clin Nutr. 1987;46(6):912-25.
103. Ceesay SM, Prentice AM, Cole TJ, Foord F, Weaver LT, Poskitt EM, et al. Effects on birth weight and perinatal mortality of maternal dietary supplements in rural Gambia: 5 year randomised controlled trial. BMJ. 1997;315(7111):786-90.
104. Khulan B, Cooper WN, Skinner BM, Bauer J, Owens S, Prentice AM, et al. Periconceptional maternal micronutrient supplementation is associated with widespread gender related changes in the epigenome: a study of a unique resource in the Gambia. Hum Mol Genet. 2012;21(9):2086-101.
105. Cooper WN, Khulan B, Owens S, Elks CE, Seidel V, Prentice AM, 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(5):1782-90.
106. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. Journal of epidemiology and community health. 2012;66(12):1182-6.
107. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: guidance for producers and users of evidence. Medical Research Council, 2012.
108. Fulford AJ, Rayco-Solon P, Prentice AM. Statistical modelling of the seasonality of preterm delivery and intrauterine growth restriction in rural Gambia. Paediatr Perinat Epidemiol. 2006;20(3):251-9.
109. Barker DJ, Eriksson JG, Forsen T, Osmond C. Fetal origins of adult disease: strength of effects and biological basis. Int J Epidemiol. 2002;31(6):1235-9.
110. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet. 2009;10(5):295-304.
111. Dominguez-Salas P, Cox SE, Prentice AM, Hennig BJ, Moore SE. Maternal nutritional status, C(1) metabolism and offspring DNA methylation: a review of current evidence in human subjects. Proc Nutr Soc. 2012;71(1):154-65.

112. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol. 2003;23(15):5293-300.
113. Koc H, Mar MH, Ranasinghe A, Swenberg JA, Zeisel SH. Quantitation of choline and its metabolites in tissues and foods by liquid chromatography/electrospray ionization-isotope dilution mass spectrometry. Anal Chem. 2002;74(18):4734-40.
114. Innis SM, Davidson AG, Melynk S, James SJ. Choline-related supplements improve abnormal plasma methionine-homocysteine metabolites and glutathione status in children with cystic fibrosis. Am J Clin Nutr. 2007;85(3):702-8.
115. Innis SM, Hasman D. Evidence of choline depletion and reduced betaine and dimethylglycine with increased homocysteine in plasma of children with cystic fibrosis. J Nutr. 2006;136(8):2226-31.
116. Midttun O, Hustad S, Solheim E, Schneede J, Ueland PM. Multianalyte quantification of vitamin B6 and B2 species in the nanomolar range in human plasma by liquid chromatography-tandem mass spectrometry. Clinical chemistry. 2005;51(7):1206-16.
117. Vuilleumier JP, Keller HE, Keck E. Clinical chemical methods for the routine assessment of the vitamin status in human populations. Part III: The apoenzyme stimulation tests for vitamin B1, B2 and B6 adapted to the Cobas-Bio analyzer. International journal for vitamin and nutrition research Internationale Zeitschrift fur Vitamin- und Ernahrungsforschung Journal international de vitaminologie et de nutrition. 1990;60(2):126-35.
118. Gibson R. Principles of Nutritional Assessment. 2nd ed. New York: Oxford University Press, Inc.; 2005.
119. WHO/FAO. Vitamin and mineral requirements in human nutrition. 2nd ed. Rome 2004.
120. Dietary reference intakes for folate, thiamine, riboflavin, niacin, vitamin B12, panthothenic acid, biotine, and choline. Institute of Medicine, National Academy of Sciences, 1998.
121. Dhonukshe-Rutten RA, de Vries JH, de Bree A, van der Put N, van Staveren WA, de Groot LC. Dietary intake and status of folate and vitamin B12 and their association with homocysteine and cardiovascular disease in European populations. Eur J Clin Nutr. 2009;63(1):18-30.
122. Hill MH, 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(2):273-8.
123. Green R. Indicators for assessing folate and vitamin B-12 status and for monitoring the efficacy of intervention strategies. Am J Clin Nutr. 2011;94(2):666S-72S.
124. Loikas S, Lopponen M, Suominen P, Moller J, Irjala K, Isoaho R, et al. RIA for serum holo-transcobalamin: method evaluation in the clinical laboratory and reference interval. Clinical chemistry. 2003;49(3):455-62.

125. Refsum H, Smith AD, Ueland PM, Nexo E, Clarke R, McPartlin J, et al. Facts and recommendations about total homocysteine determinations: an expert opinion. Clinical chemistry. 2004;50(1):3-32.
126. Prentice AM, Whitehead RG, Roberts SB, Paul AA. Long-term energy balance in child-bearing Gambian women. Am J Clin Nutr. 1981;34(12):2790-9.
127. Kennedy G, Fanou-Fogny N, Seghieri C, Arimond M, Koreissi Y, Dossa R, et al. Food groups associated with a composite measure of probability of adequate intake of 11 micronutrients in the diets of women in urban Mali. J Nutr. 2010;140(11):2070S-8S.
128. Becquey E, Martin-Prevel Y. Micronutrient adequacy of women's diet in urban Burkina Faso is low. J Nutr. 2010;140(11):2079S-85S.
129. Gossell-Williams M, Fletcher H, McFarlane-Anderson N, Jacob A, Patel J, Zeisel S. Dietary intake of choline and plasma choline concentrations in pregnant women in Jamaica. West Indian Med J. 2005;54(6):355-9.
130. Jensen HH, Batres-Marquez SP, Carriquiry A, Schalinske KL. Choline in the diets of the U.S. population: NHANES 2003-2004. FASEB J. 2007(21):lb219.
131. Chiuve SE, Giovannucci EL, Hankinson SE, Zeisel SH, Dougherty LW, Willett WC, et al. The association between betaine and choline intakes and the plasma concentrations of homocysteine in women. Am J Clin Nutr. 2007;86(4):1073-81.
132. Jacob RA, Jenden DJ, Allman-Farinelli MA, Swendseid ME. Folate nutriture alters choline status of women and men fed low choline diets. J Nutr. 1999;129(3):712-7.
133. Moat SJ, Ashfield-Watt PA, Powers HJ, Newcombe RG, McDowell IF. Effect of riboflavin status on the homocysteine-lowering effect of folate in relation to the MTHFR (C677T) genotype. Clinical chemistry. 2003;49(2):295-302.
134. Garry PJ, Owen GM. An automated flavin adenine dinucleotide-dependent glutathione reductase assay for assessing riboflavin nutriture. Am J Clin Nutr. 1976;29(6):663-74.
135. Thurnham DI, Rathakette P. Incubation of NAD(P)H2: glutathione oxidoreductase (EC 1.6.4.2) with flavin adenine dinucleotide for maximal stimulation in the measurement of riboflavin status. Br J Nutr. 1982;48(3):459-66.
136. Fischer LM, da Costa KA, Kwock L, Galanko J, Zeisel SH. Dietary choline requirements of women: effects of estrogen and genetic variation. Am J Clin Nutr. 2010;92(5):1113-9.
137. Niculescu MD, Zeisel SH. Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. J Nutr. 2002;132(8 Suppl):2333S-5S.
138. Yajnik CS, Deshpande SS, Jackson AA, Refsum H, Rao S, Fisher DJ, et al. Vitamin B12 and folate concentrations during pregnancy and insulin resistance in the offspring: the Pune Maternal Nutrition Study. Diabetologia. 2008;51(1):29-38.
139. Hudson GJ. Food intake in a west African village. Estimation of food intake from a shared bowl. Br J Nutr. 1995;73(4):551-69.

140. Block G, Hartman AM. Issues in reproducibility and validity of dietary studies. Am J Clin Nutr. 1989;50(5 Suppl):1133-8; discussion 231-5.
141. Zeisel SH, Mar MH, Howe JC, Holden JM. Concentrations of choline-containing compounds and betaine in common foods. J Nutr. 2003;133(5):1302-7.
142. Barikmo I. Food composition table for Mali. Akershus University College, Norway, 2004.
143. Widdowson M. The composition of foods: Food Standards Agency; 2002.
144. Huybregts LF, Roberfroid DA, Kolsteren PW, Van Camp JH. Dietary behaviour, food and nutrient intake of pregnant women in a rural community in Burkina Faso. Maternal & child nutrition. 2009;5(3):211-22.
145. Velzing-Aarts FV, Holm PI, Fokkema MR, van der Dijs FP, Ueland PM, Muskiet FA. Plasma choline and betaine and their relation to plasma homocysteine in normal pregnancy. Am J Clin Nutr. 2005;81(6):1383-9.
146. Koutmos M, Kabil O, Smith JL, Banerjee R. Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine {beta}-synthase. Proc Natl Acad Sci U S A. 2010;107(49):20958-63.
147. Institute of Medicine, National Academy of Sciences USA. Choline. Dietary reference intakes for folate, thiamin, riboflavin, niacin, vitamin B12, panthothenic acid, biotin, and choline. Washington D.C.: National Academy Press; 1998. p. 390-422.
148. Cho E, Willett WC, Colditz GA, Fuchs CS, Wu K, Chan AT, et al. Dietary choline and betaine and the risk of distal colorectal adenoma in women. Journal of the National Cancer Institute. 2007;99(16):1224-31.
149. Olsen A, Halkjaer J, van Gils CH, Buijsse B, Verhagen H, Jenab M, et al. Dietary intake of the water-soluble vitamins B1, B2, B6, B12 and C in 10 countries in the European Prospective Investigation into Cancer and Nutrition. Eur J Clin Nutr. 2009;63 Suppl 4:S122-49.
150. Sassine A. Assessment of the effect of Ramadan on dietary intake and nutritional status of adult women in West Kiang, The Gambia. MSc summer project: LSHTM, 2011.
151. Hudson GJ. Food intake in a West African village 1992.
152. Heini A, Schutz Y, Jequier E. Twenty-four-hour energy expenditure in pregnant and nonpregnant Gambian women, measured in a whole-body indirect calorimeter. Am J Clin Nutr. 1992;55(6):1078-85.
153. Singh J, Prentice AM, Diaz E, Coward WA, Ashford J, Sawyer M, et al. Energy expenditure of Gambian women during peak agricultural activity measured by the doubly-labelled water method. Br J Nutr. 1989;62(2):315-29.
154. Food and Drug Administration. Available from: http://www.accessdata.fda.gov/cdrh_docs/pdf/K965186.pdf.

155. Shaw GM, Finnell RH, Blom HJ, Carmichael SL, Vollset SE, Yang W, et al. Choline and risk of neural tube defects in a folate-fortified population. Epidemiology. 2009;20(5):714-9.
156. Bingham SA. Limitations of the various methods for collecting dietary intake data. Annals of nutrition & metabolism. 1991;35(3):117-27.
157. Poirier LA, Wise CK, Delongchamp RR, Sinha R. Blood determinations of S-adenosylmethionine, S-adenosylhomocysteine, and homocysteine: correlations with diet. Cancer Epidemiol Biomarkers Prev. 2001;10(6):649-55.
158. Cho E, Zeisel SH, Jacques P, Selhub J, Dougherty L, Colditz GA, et al. 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. 2006;83(4):905-11.
159. Dalmeijer GW, Olthof MR, Verhoef P, Bots ML, van der Schouw YT. Prospective study on dietary intakes of folate, betaine, and choline and cardiovascular disease risk in women. Eur J Clin Nutr. 2008;62(3):386-94.
160. Dominguez-Salas P. Assessment of dietary methyl donor load of women in reproductive age in Keneba, The Gambia. MSc Public Health Nutrition: LSHTM; 2007.
161. Choi SW, Friso S. Epigenetics: A New Bridge between Nutrition and Health. Adv Nutr. 2010;1(1):8-16.
162. Taberlay PC, Jones PA. DNA methylation and cancer. Progress in drug research Fortschritte der Arzneimittelforschung Progres des recherches pharmaceutiques. 2011;67:1-23.
163. Villeneuve LM, Reddy MA, Natarajan R. Epigenetics: deciphering its role in diabetes and its chronic complications. Clinical and experimental pharmacology & physiology. 2011;38(7):401-9.
164. Zeisel SH. Importance of methyl donors during reproduction. Am J Clin Nutr. 2009;89(2):673S-7S.
165. Bhaskar LV, Murthy J, Venkatesh Babu G. Polymorphisms in genes involved in folate metabolism and orofacial clefts. Arch Oral Biol. 2011;56(8):723-37.
166. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA, 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(11):e7845.
167. Dolinoy DC, Das R, Weidman JR, Jirtle RL. Metastable epialleles, imprinting, and the fetal origins of adult diseases. Pediatr Res. 2007;61(5 Pt 2):30R-7R.
168. Molloy AM, Mills JL, Cox C, Daly SF, Conley M, Brody LC, et al. Choline and homocysteine interrelations in umbilical cord and maternal plasma at delivery. Am J Clin Nutr. 2005;82(4):836-42.

169. Friesen RW, Novak EM, Hasman D, Innis SM. Relationship of dimethylglycine, choline, and betaine with oxoproline in plasma of pregnant women and their newborn infants. J Nutr. 2007;137(12):2641-6.
170. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215.
171. Laboratories B-R. Chelex® 100 and Chelex 20 Chelating Ion Exchange Resin Instruction Manual. BIO-RAD; 2000.
172. Shen L, Guo Y, Chen X, Ahmed S, Issa JP. Optimizing annealing temperature overcomes bias in bisulfite PCR methylation analysis. BioTechniques. 2007;42(1):48, 50, 2 passim.
173. WHO DoNfHaD. WHO Child Growth Standards. Lenght/height-for-age, weight-for-age, weight-for-lenght, weigth-for-heaight and body mass index-for-age. Methods and development. 2006.
174. Murphy MM, Fernandez-Ballart JD. Homocysteine in pregnancy. Advances in clinical chemistry. 2011;53:105-37.
175. van Driel LM, Eijkemans MJ, de Jonge R, de Vries JH, van Meurs JB, Steegers EA, et al. Body mass index is an important determinant of methylation biomarkers in women of reproductive ages. J Nutr. 2009;139(12):2315-21.
176. Feng Y, Xue WJ, Li P, Sha ZY, Huang H, Rui L, et al. RASSF1A Hypermethylation is Associated with Aflatioxin B1 and Polycyclic Aromatic Hydrocarbon Exposure in Hepatocellular Carcinoma. Hepato-gastroenterology. 2012;59(118):1883-8.
177. Herceg Z, Paliwal A. Epigenetic mechanisms in hepatocellular carcinoma: how environmental factors influence the epigenome. Mutat Res. 2011;727(3):55-61.
178. Bates CJ, Fuller NJ, Prentice AM. Folate status during pregnancy and lactation in a West African rural community. Human nutrition Clinical nutrition. 1986;40(1):3-13.
179. Reddy VA, Bates CJ, Goh SG, Rowland MG, Greenwood AM, Greenwood B, et al. Riboflavin, folate and vitamin C status of Gambian women during pregnancy: a comparison between urban and rural communities. Trans R Soc Trop Med Hyg. 1987;81(6):1033-7.
180. Prentice AM, Prentice A, Lamb WH, Lunn PG, Austin S. Metabolic consequences of fasting during Ramadan in pregnant and lactating women. Human nutrition Clinical nutrition. 1983;37(4):283-94.
181. Gemma C, Sookoian S, Alvarinas J, Garcia SI, Quintana L, Kanevsky D, et al. Maternal pregestational BMI is associated with methylation of the PPARGC1A promoter in newborns. Obesity (Silver Spring). 2009;17(5):1032-9.
182. Carranza D, Van Vliet G, Polak M. [Congenital hypothyroidism]. Annales d'endocrinologie. 2006;67(4):295-302.

183. Yi P, Melnyk S, Pogribna M, Pogribny IP, Hine RJ, James SJ. Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation. J Biol Chem. 2000;275(38):29318-23.
184. Herbstman JB, Tang D, Zhu D, Qu L, Sjodin A, Li Z, et al. Prenatal exposure to polycyclic aromatic hydrocarbons, benzo[a]pyrene-DNA adducts, and genomic DNA methylation in cord blood. Environ Health Perspect. 2012;120(5):733-8.
185. Perera F, Tang D, Whyatt R, Lederman SA, Jedrychowski W. DNA damage from polycyclic aromatic hydrocarbons measured by benzo[a]pyrene-DNA adducts in mothers and newborns from Northern Manhattan, the World Trade Center Area, Poland, and China. Cancer Epidemiol Biomarkers Prev. 2005;14(3):709-14.
186. Macchia PE, Lapi P, Krude H, Pirro MT, Missero C, Chiovato L, et al. PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis. Nat Genet. 1998;19(1):83-6.
187. Stokes E, Dumbaya I, Owens S, Brabin L. The right to remain silent: a qualitative study of the medical and social ramifications of pregnancy disclosure for Gambian women. BJOG. 2008;115(13):1641-7; discussion 7.
188. Zeisel SH. Choline: critical role during fetal development and dietary requirements in adults. Annu Rev Nutr. 2006;26:229-50.
189. van Straten EM, Bloks VW, Huijkman NC, Baller JF, van Meer H, Lutjohann D, et al. The liver X-receptor gene promoter is hypermethylated in a mouse model of prenatal protein restriction. American journal of physiology Regulatory, integrative and comparative physiology. 2010;298(2):R275-82.
190. Hoffman DR, Cornatzer WE, Duerre JA. Relationship between tissue levels of S-adenosylmethionine, S-adenylhomocysteine, and transmethylation reactions. Canadian journal of biochemistry. 1979;57(1):56-65.
191. van Ewijk R. Long-term health effects on the next generation of Ramadan fasting during pregnancy. Journal of health economics. 2011;30(6):1246-60.
192. Robinson T, Raisler J. "Each one is a doctor for herself": Ramadan fasting among pregnant Muslim women in the United States. Ethnicity & disease. 2005;15(1 Suppl 1):S1-99-103.
193. Ozturk E, Balat O, Ugur MG, Yazicioglu C, Pence S, Erel O, et al. Effect of Ramadan fasting on maternal oxidative stress during the second trimester: a preliminary study. The journal of obstetrics and gynaecology research. 2011;37(7):729-33.
194. Almond D, Mazumder B. Health Capital and the Prenatal Environment: The Effect of Ramadan Observance During Pregnancy. Am Econ J-Appl Econ. 2011;3(4):56-85.
195. El-Maarri O, Becker T, Junen J, Manzoor SS, Diaz-Lacava A, Schwaab R, et al. Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. Hum Genet. 2007;122(5):505-14.

196. Alwasel SH, Abotalib Z, Aljarallah JS, Osmond C, Alkharaz SM, Alhazza IM, et al. Sex differences in birth size and intergenerational effects of intrauterine exposure to Ramadan in Saudi Arabia. American journal of human biology : the official journal of the Human Biology Council. 2011;23(5):651-4.
197. Adkins RM, Thomas F, Tylavsky FA, Krushkal J. Parental ages and levels of DNA methylation in the newborn are correlated. BMC medical genetics. 2011;12:47.
198. Adkins RM, Krushkal J, Tylavsky FA, Thomas F. Racial differences in gene-specific DNA methylation levels are present at birth. Birth Defects Res A Clin Mol Teratol. 2011;91(8):728-36.
199. Crider KS, Yang TP, Berry RJ, Bailey LB. Folate and DNA methylation: a review of molecular mechanisms and the evidence for folate's role. Adv Nutr. 2012;3(1):21-38.
200. Kim YI. Folate and DNA methylation: a mechanistic link between folate deficiency and colorectal cancer? Cancer Epidemiol Biomarkers Prev. 2004;13(4):511-9.
201. Sinclair KD, Allegrucci C, Singh R, Gardner DS, Sebastian S, Bispham J, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. Proc Natl Acad Sci U S A. 2007;104(49):19351-6.
202. Waterland RA, Dolinoy DC, Lin JR, Smith CA, Shi X, Tahiliani KG. Maternal methyl supplements increase offspring DNA methylation at Axin Fused. Genesis. 2006;44(9):401-6.
203. Fryer AA, Nafee TM, Ismail KM, Carroll WD, Emes RD, Farrell WE. LINE-1 DNA methylation is inversely correlated with cord plasma homocysteine in man: a preliminary study. Epigenetics. 2009;4(6):394-8.
204. Eaton BM, Sooranna SR. Regulation of the choline transport system in superfused microcarrier cultures of BeWo cells. Placenta. 1998;19(8):663-9.
205. Grassl SM. Choline transport in human placental brush-border membrane vesicles. Biochim Biophys Acta. 1994;1194(1):203-13.
206. van der Aa EM, Wouterse AC, Peereboom-Stegeman JH, Russel FG. Uptake of choline into syncytial microvillus membrane vesicles of human term placenta. Biochem Pharmacol. 1994;47(3):453-6.
207. 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 and childbirth. 2012;12:107.
208. McKay JA, Mathers JC. Diet induced epigenetic changes and their implications for health. Acta Physiol (Oxf). 2011;202(2):103-18.
209. Paul AA, Bates CJ, Prentice A, Day KC, Tsuchiya H. Zinc and phytate intake of rural Gambian infants, contflbutions from breastmilk and weaning foods. International Journal of Food Sciences and Nurrition. 1998(49):141-55.

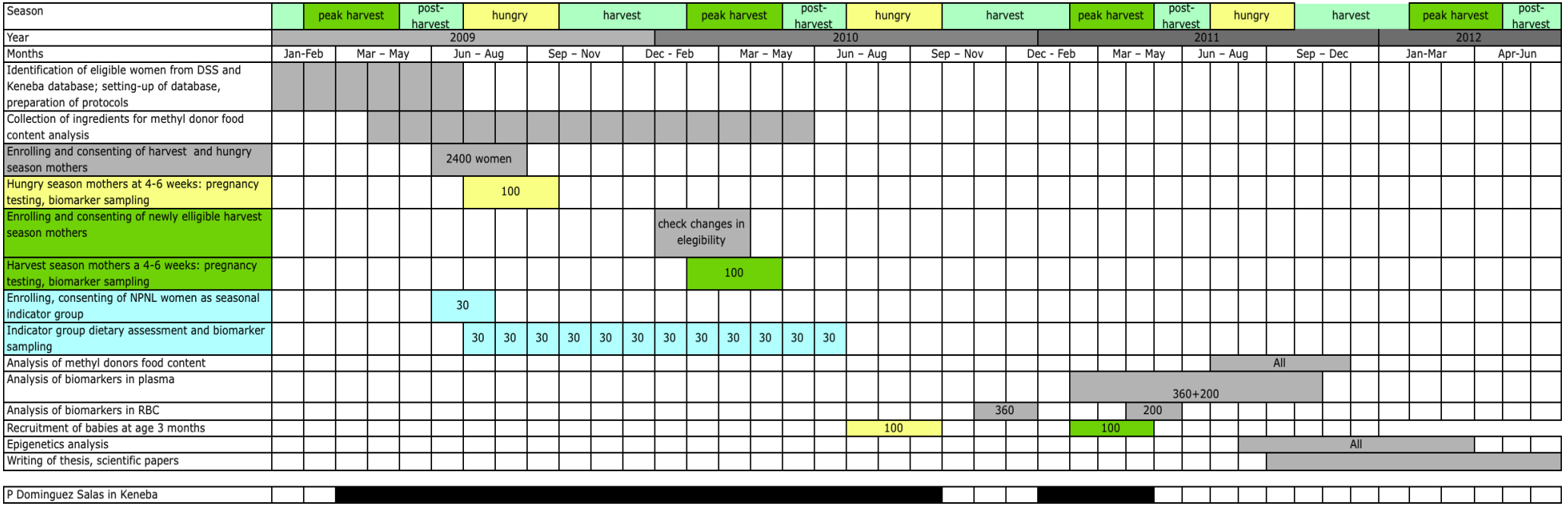
210. Zhang YJ, Ahsan H, Chen Y, Lunn RM, Wang LY, Chen SY, et al. High frequency of promoter hypermethylation of RASSF1A and p16 and its relationship to aflatoxin B1-DNA adduct levels in human hepatocellular carcinoma. Molecular carcinogenesis. 2002;35(2):85-92.
211. Turner PC, Moore SE, Hall AJ, Prentice AM, Wild CP. Modification of immune function through exposure to dietary aflatoxin in Gambian children. Environ Health Perspect. 2003;111(2):217-20.
212. Pfeifer GP, Dammann R. Methylation of the tumor suppressor gene RASSF1A in human tumors. Biochemistry Biokhimiia. 2005;70(5):576-83.
213. Zhang YJ, Chen Y, Ahsan H, Lunn RM, Lee PH, Chen CJ, et al. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation and its relationship to aflatoxin B1-DNA adducts and p53 mutation in hepatocellular carcinoma. Int J Cancer. 2003;103(4):440-4.
214. Rampersaud GC, Kauwell GP, Hutson AD, Cerda JJ, Bailey LB. Genomic DNA methylation decreases in response to moderate folate depletion in elderly women. Am J Clin Nutr. 2000;72(4):998-1003.
215. van den Donk M, van Engeland M, Pellis L, Witteman BJ, Kok FJ, Keijer J, et al. Dietary folate intake in combination with MTHFR C677T genotype and promoter methylation of tumor suppressor and DNA repair genes in sporadic colorectal adenomas. Cancer Epidemiol Biomarkers Prev. 2007;16(2):327-33.
216. Portela A, Esteller M. Epigenetic modifications and human disease. Nat Biotechnol. 2010;28(10):1057-68.
217. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462(7271):315-22.
218. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet. 2009;41(2):178-86.
219. Jiang X, Yan J, West AA, Perry CA, Malysheva OV, Devapatla S, et al. Maternal choline intake alters the epigenetic state of fetal cortisol-regulating genes in humans. FASEB J. 2012.
220. Jackson AA, Robinson SM. Dietary guidelines for pregnancy: a review of current evidence. Public Health Nutr. 2001;4(2B):625-30.
221. Mudd SH, Brosnan JT, Brosnan ME, Jacobs RL, Stabler SP, Allen RH, et al. Methyl balance and transmethylation fluxes in humans. Am J Clin Nutr. 2007;85(1):19-25.
222. Baxmann AC, Ahmed MS, Marques NC, Menon VB, Pereira AB, Kirsztajn GM, et al. Influence of muscle mass and physical activity on serum and urinary creatinine and serum cystatin C. Clin J Am Soc Nephrol. 2008;3(2):348-54.

223. Nijhout HF, Reed MC, Ulrich CM. Mathematical models of folate-mediated one-carbon metabolism. Vitamins and hormones. 2008;79:45-82.
224. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. Trends Genet. 2002;18(7):348-51.
225. Foley JW, Katagiri F. Unsupervised reduction of random noise in complex data by a row-specific, sorted principal component-guided method. BMC bioinformatics. 2008;9:508.
226. Miyaki K. Genetic polymorphisms in homocysteine metabolism and response to folate intake: a comprehensive strategy to elucidate useful genetic information. Journal of epidemiology / Japan Epidemiological Association. 2010;20(4):266-70.
227. Fischer LM, daCosta KA, Kwock L, Stewart PW, Lu TS, Stabler SP, et al. Sex and menopausal status influence human dietary requirements for the nutrient choline. Am J Clin Nutr. 2007;85(5):1275-85.
228. Ray JG, Vermeulen MJ, Boss SC, Cole DE. Increased red cell folate concentrations in women of reproductive age after Canadian folic acid food fortification. Epidemiology. 2002;13(2):238-40.
229. Ehrlich M. DNA methylation in cancer: too much, but also too little. Oncogene. 2002;21(35):5400-13.
230. Wild CP, Hudson GJ, Sabbioni G, Chapot B, Hall AJ, Wogan GN, et al. Dietary intake of aflatoxins and the level of albumin-bound aflatoxin in peripheral blood in The Gambia, West Africa. Cancer Epidemiol Biomarkers Prev. 1992;1(3):229-34.
231. Wild CP, Yin F, Turner PC, Chemin I, Chapot B, Mendy M, et al. Environmental and genetic determinants of aflatoxin-albumin adducts in the Gambia. Int J Cancer. 2000;86(1):1-7.
232. Constancia M, Angiolini E, Sandovici I, Smith P, Smith R, Kelsey G, et al. Adaptation of nutrient supply to fetal demand in the mouse involves interaction between the Igf2 gene and placental transporter systems. Proc Natl Acad Sci U S A. 2005;102(52):19219-24.
233. Nijhout HF, Reed MC, Lam SL, Shane B, Gregory JF, 3rd, Ulrich CM. In silico experimentation with a model of hepatic mitochondrial folate metabolism. Theoretical biology & medical modelling. 2006;3:40.
234. Lillycrop KA, Phillips ES, Jackson AA, Hanson MA, Burdge GC. Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. J Nutr. 2005;135(6):1382-6.
235. Davison JM, Mellott TJ, Kovacheva VP, Blusztajn JK. Gestational choline supply regulates methylation of histone H3, expression of histone methyltransferases G9a (Kmt1c) and Suv39h1 (Kmt1a), and DNA methylation of their genes in rat fetal liver and brain. J Biol Chem. 2009;284(4):1982-9.

236. Kovacheva VP, Mellott TJ, Davison JM, Wagner N, Lopez-Coviella I, Schnitzler AC, et al. Gestational choline deficiency causes global and Igf2 gene DNA hypermethylation by up-regulation of Dnmt1 expression. J Biol Chem. 2007;282(43):31777-88.
237. Haider BA, Bhutta ZA. Multiple-micronutrient supplementation for women during pregnancy. Cochrane Database Syst Rev. 2006(4):CD004905.
238. Resseguie ME, da Costa KA, Galanko JA, Patel M, Davis IJ, Zeisel SH. Aberrant estrogen regulation of PEMT results in choline deficiency-associated liver dysfunction. J Biol Chem. 2011;286(2):1649-58.
239. Zeisel SH. Diet-gene interactions underlie metabolic individuality and influence brain development: implications for clinical practice derived from studies on choline metabolism. Annals of nutrition & metabolism. 2012;60 Suppl 3:19-25.
240. Glenn MJ, Kirby ED, Gibson EM, Wong-Goodrich SJ, Mellott TJ, Blusztajn JK, et al. Age-related declines in exploratory behavior and markers of hippocampal plasticity are attenuated by prenatal choline supplementation in rats. Brain Res. 2008;1237:110-23.
241. Niculescu MD, Craciunescu CN, Zeisel SH. Dietary choline deficiency alters global and gene-specific DNA methylation in the developing hippocampus of mouse fetal brains. FASEB J. 2006;20(1):43-9.
242. Niculescu MD, Craciunescu CN, Zeisel SH. Gene expression profiling of choline-deprived neural precursor cells isolated from mouse brain. Brain Res Mol Brain Res. 2005;134(2):309-22.
243. Meck WH, Williams CL. Metabolic imprinting of choline by its availability during gestation: implications for memory and attentional processing across the lifespan. Neurosci Biobehav Rev. 2003;27(4):385-99.
244. da Costa KA, Kozyreva OG, Song J, Galanko JA, Fischer LM, Zeisel SH. Common genetic polymorphisms affect the human requirement for the nutrient choline. FASEB J. 2006;20(9):1336-44.
245. Niculescu MD, Craciunescu CN, Zeisel SH. Gene expression profiling of choline-deprived neural precursor cells isolated from mouse brain. Brain Res 2005;134(2):309-22.
246. Mehedint MG, Niculescu MD, Craciunescu CN, Zeisel SH. Choline deficiency alters global histone methylation and epigenetic marking at the Re1 site of the calbindin 1 gene. FASEB J. 2010;24(1):184-95.
247. Smallwood SA, Kelsey G. De novo DNA methylation: a germ cell perspective. Trends Genet. 2012;28(1):33-42.
248. SACN. The influence of maternal, fetal and child nutrition on the development of chronic disease in later life. 2011.
249. Abbott A. Europe to map the human epigenome. Nature. 2011;477(7366):518.

APPENDICES

APPENDIX I: Study timeline



APPENDIX II: Evidence of copyright retention

Licence to Publish

*In order to publish your article the **Proceedings of the Nutrition Society** needs your agreement in writing. Please read the terms of this licence, sign the form, and send it as soon as your paper is accepted for publication (the form may be photocopied)*

Title of article (and reference no): Maternal nutritional status, one-carbon metabolism and offspring DNA methylation: A review of current evidence in humans - PNS-2011-000191R

……

Name(s) of Author(s): Paula Dominguez-Salas, Sharon Cox, Andrew Prentice, Branwen Hennig, and Sophie Moore

……

By signing this form;

- **You (the author/s or other copyright owner) agree** to grant the Nutrition Society the exclusive right to reproduce and to distribute your article (including the abstract) throughout the world in any language in printed, electronic or any other medium, and in turn to authorise others (including Reproduction Rights Organisations such as the Copyright Licensing Agency and the Copyright Clearance Center) to do the same. You agree that we may publish your article in *Proceedings of the Nutrition Society*, and that we may sell or distribute it within the journal, on its own, or with other related material.

- **You warrant** that the article is your original work, has not previously been published, and is not currently under consideration by another journal. If it contains material which is someone else's copyright, you warrant that you have obtained the permission of the copyright owner and that the material is clearly identified and acknowledged within the text. You also **warrant** that the article does not, to the best of your knowledge, contain anything which is libellous, illegal, or infringes anyone's copyright or other rights.

You **warrant** that the agreement of those whom you acknowledge in the article has been obtained. The source of funding should be identified in the acknowledgement section of the manuscript. All potential conflicts of interest, or financial interests of the author in a product or company which is relevant to the article, should be declared at the end of this form.

- **You assert** your Moral Right to be identified as the author(s), and **we affirm** that we will respect your rights as the author(s). That is, we will ensure that your name(s) is/are always clearly associated with the article and, while you do allow us to make necessary editorial changes, we will not make any substantial alteration without consulting you. We will have your article fairly reviewed by impartial referees, and will let you know the outcome as soon as possible. When the article is accepted, we undertake to publish it within 9 months. If it is rejected, or we fail to publish it within 9 months, this agreement is automatically cancelled and all the rights come back to you. On publication, we will send you a PDF of your article. We will do everything we reasonably can to maximise the visibility of the journal, and of your article within it.

- **Copyright remains yours (or that of your Institution)**, and this we will acknowledge. However, you authorise us to act on your behalf to defend your copyright if anyone should infringe it, and to retain half of any damages awarded, after deducting our costs. You retain the right to use your own article (provided you acknowledge the published original in standard bibliographic citation form) as follows, as long as you do not sell it in ways which would conflict directly with our commercial interests: for the internal educational or other purposes of your own institution or organisation; mounted on your own or your institution's website; posted to free public servers of preprints and/or articles in your subject area; in whole or in part, as the basis for your own further publications or spoken presentations, although ideally not before 6 months have elapsed after publication.

- **You warrant** that you have no potential conflict of interest pertaining to your contribution to the Journal. If you have such a conflict of interest, please declare it now.

  Potential conflict of interest

  ……

1 To be completed if copyright belongs to you

I hereby assign to The Nutrition Society licence to publish in all forms and media in the said contribution, including in any supplementary materials that I/we may author in support of the online version.

Signed (tick one) □ the sole author(s)

□ one author authorised to execute this transfer on behalf of all the authors of the above article

Name (block letters).....

Institution/Company .....

Signature:.....Date:..... .....

2 US Government exemption

I/we certify that the paper above was written in the course of employment by the United States Government so that no copyright exists.

Signature:.....Name (Block letters): .....

Please make a duplicate of this form for your own records

Please return this form by post or fax to: Publications Office, *Proceedings of the Nutrition Society*, The Nutrition Society, 10 Cambridge Court, 210 Shepherds Bush Road, London W6 7NJ, UK [fax +44 (0)20 7602 1756].

APPENDIX III: Pictures of fieldwork in The Gambia

A) Field station

B) Taring method (indicator group)

Panel A: Staple (rice)

Panel B: Sauce (vegetal oil stew)

Panel C: Vegetables and/or fish

APPENDIX IV: Information sheet and consent form for main group

INFORMATION SHEET FOR WOMEN OF MAIN STUDY
Medical Research Council Laboratories

Epigenesis in humans: Can maternal methyl-donor-deficient diets induce epigenetic alterations in their offspring?
(SCC 1151)

Subject Information Sheet (Main study participants)
Version number 02 20th April, 2009
(If necessary, to be read to participants in their own language)

What is the purpose of this study?

The diet of many women in West Kiang is often poor. This sometimes leads to problems of poor nutritional status during pregnancy, and may also affect the health of their babies when they are adult. This is because the mother's nutrition during the first months of pregnancy is very important for the development of her infant. Since the diet in West Kiang is not the same during the different seasons, we would like to look in more detail at whether the differences in the intake of some specific substances such as vitamins, reflects on the mothers level of nutrition and then leads to any differences in the babies. We would like to invite you to participate.

Before you decide to participate it is important for you to understand why the research is being done and what it will involve. Please take time to listen to this carefully and discuss it with others if you wish. Please ask if there is anything which is unclear or if you would like more information. Take time to decide whether or not you wish to take part.

What does participation in the study mean for you?

We are trying to find out whether the diet during the first stages pregnancy is rich in certain substances such as vitamins. We also want to find out if this leads to internal differences in the babies. To do this, we would like to enrol you in this study and if you become pregnant, we would like to conduct a series of tests on you and then on your infant when he/she is born.

The following will be required if you agree to take part:

Once you have consented, you will be asked a short questionnaire on the date of your last period and you will be visited every month by a member of the project team with the same short questionnaire. When you report a missing period, we would like to collect a small blood sample from one of your veins so that we can measure the amount of certain nutrients in your blood. The amount we will collect is very small (10mls in total, equivalent to two teaspoons). When you report two missing periods, we will ask you to provide a small urine sample, which we will used to test for pregnancy. You will then be visited by the midwife to confirm the result of this test and, if you are pregnant, you will be enrolled in the study. If you are not pregnant, you will continue to be visited by the field worker. If you become pregnant during the time of the study, we will collect samples at two further timepoints:

1. When you are in labour, an MRC field assistant living in or close to your village will be informed. Following delivery, and if we have your permission, he/she will collect your placenta from the midwife/traditional birth attendant and will use this to take samples of cord blood and placental material.

2. When your infant is 12 weeks of age, we will invite you and your infant to Keneba for some further measurements on your infant. We will take a small blood sample from one of their veins. The amount we will collect is very small (3mls in total, equivalent to 1 teaspoon). We will also need to take a buccal swab (gently rubbing the inner cheek with cotton), which is painless for your baby. We will collect a sample of his/her faeces and we will ask you to pluck some hairs (30) from his/her head. All these samples will be equally used to extract DNA.

All samples collected from you and your baby will be processed in the laboratory in Keneba for storage. Your blood sample will be used to measure nutrient levels and to extract DNA so that we can look for genes such as those that are related to the possible future diseases. The samples collected from you baby will be used to measure whether the genes your baby is born with are affected by what you eat and what your nutritional status is.

All the samples will be transferred and analysed in laboratories in the UK, US and Canada for. This is because we do not have the equipment in The Gambia for measuring the factors we are interested in. All of the information obtained from abroad will then be sent back to the investigators in The Gambia.

All information which is collected during the course of this study will be kept strictly confidential and you will only be identified by an ID number.

It is up to you to decide whether or not to take part. If you do decide to take part you will be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. If you have any questions, please ask. We hope that you will agree to participate.

Thank you very much.

Principal investigator:
Paula Dominguez Salas & Dr Branwen Hennig on behalf of the Epigenesis study team
Mobile number for enquiries: 9391334

CONSENT FORM FOR WOMEN OF MAIN STUDY

Medical Research Council Laboratories

Epigenesis in humans: Can maternal methyl-donor-deficient diets induce epigenetic alterations in their offspring?

(SCC 1151)

Consent Form (Main study participants):

Version number 02 20th April, 2009

The information sheet has been read to me and I understand it / I have read and understood the information sheet.

I understand what participation in the study means for me and my infant.

I understand that the information regarding me that is collected in the course of this study will remain confidential.

I understand that laboratory tests will be done on the blood and placental samples from me as well as on the blood, hair, faeces and buccal sample taken of my baby.

I understand that the samples will be sent to the UK, US or Canada for laboratory testing. I also understand that part of the blood samples collected will be stored for future genetic analyses conducted by the investigators running this study.

I understand that if either myself or my infant gets sick during the study period, I can go to the clinic where study staff are providing care, and that we will be examined and treated for free.

I understand that I am free to take part in the study or refuse, and that I can withdraw either myself of my infant from the study at any time, and without giving any reason. Deciding not to take part or to withdraw from the study will not affect the care that I or any of my family is normally entitled to.

I have had a chance to ask questions and have them answered.

Signature or thumb print of volunteer: _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

This form has been read by / I have read the above to _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

(write name of volunteer)

in a language that she understands. I believe that she has understood what I explained and that she has freely agreed to take part in the study.

Signature of field worker: _

Name of field worker: _

Date: |__|__| / |__|__| / |__|__|__|__|

APPENDIX V: Information sheet and consent form for indicator group

INFORMATION SHEET FOR WOMEN OF INDICATOR GROUP

Medical Research Council Laboratories

Epigenesis in humans: Can maternal methyl-donor-deficient diets induce epigenetic alterations in their offspring?
(SCC 1151)

Subject Information Sheet (Indicator group participants)

Version number 02 20th April, 2009

(If necessary, to be read to participants in their own language)

What is the purpose of this study?

The diet of many women in West Kiang is often poor. This sometimes leads to problems affecting health. Since the diet in West Kiang is not the same during the different seasons, we would like to look in more detail at whether the differences in the intake of some specific substances such as vitamins, reflects on the women's level of nutrition. We would like to invite you to participate.

Before you decide to participate it is important for you to understand why the research is being done and what it will involve. Please take time to listen to this carefully and discuss it with others if you wish. Please ask if there is anything which is unclear or if you would like more information. Take time to decide whether or not you wish to take part.

What does participation in the study mean for you?

We are trying to find out whether the nutritional status of women in West Kiang is adequate in certain substances such as vitamins, along the year. To do this, we would like to enrol you in this study to conduct a series of tests and measures on you.

The following will be required if you agree to take part:

Once enrolled, you will be visited at your household once a month by a fieldworker, so you will be visited a total of 12 timed during the year. The field worker will observe the meal preparation in the household, and will measure your food intake. This will happen on two days (one after the other). It is important that the fieldworker records everything that you

eat and drink during each 24 hours, and so they will need to spend the whole day with you from when you wake up in the morning until when you go to bed.

It is also important that you do not change your behaviour and you eat the same that you would eat if the field worker was not there, because what we need to know is what your normal diet is. We may ask you to give us a small amount of specific foods (about a handful) if we would like to test the amount of vitamins and other substances in this food or to give us information about the origin of the ingredients you used.

On the third day, we would like to collect a blood sample from one of your veins so that we can measure the amount of certain nutrients in your blood. The amount we will collect is very small (10mls in total, equivalent to two teaspoons).

You will also be asked a short questionnaire on the date of your last period. If you become pregnant during the time of the study, you will have to withdraw the study group. If you wish so, you can then join the group of pregnant women, whose details will be explained to you accordingly.

The samples collected from you will be transported back to the laboratory in Keneba for storage. Your blood sample will be used to measure nutrient levels and to extract DNA so that we can look for genes such as those that are related to the possible future diseases. All the samples will be transferred and analysed in laboratories in the UK, US and Canada for. This is because we do not have the equipment in The Gambia for measuring the factors we are interested in. All of the information obtained from abroad will then be sent back to the investigators in The Gambia.

All information which is collected during the course of this study will be kept strictly confidential and you will only be identified by an ID number.

It is up to you to decide whether or not to take part. If you do decide to take part you will be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. If you have any questions, please ask. We hope that you will agree to participate.

Thank you very much.

Principal investigator:
Paula Dominguez Salas & Dr Branwen Hennig on behalf of the Epigenesis study team
Mobile number for enquiries: 9391334

CONSENT FORM FOR WOMEN OF INDICATOR GROUP
Medical Research Council Laboratories

Epigenesis in humans: Can maternal methyl-donor-deficient diets induce epigenetic alterations in their offspring?

(SCC 1151)

Consent Form (Indicator group participants):

Version number 02 20th April, 2009

The information sheet has been read to me and I understand it / I have read and understood the information sheet. I understand what participation in the study means for me.

I understand that the information regarding me that is collected in the course of this study will remain confidential.

I understand that during the time of the study, I will be visited every month to repeat the same procedure (12 times in total).

I understand that I will receive fieldworkers at home who will observe the meal preparation and will measure my daily intake. I understand that he/she needs to record everything that I eat during those 2 days and so they will spend most of the day with me.

I understand that laboratory tests will be done on the blood collected from. I understand that the samples will be sent to the UK, US or Canada for laboratory testing. I also understand that part of the blood samples collected will be stored for future genetic analyses conducted by the investigators running this study.

I understand that if I get sick during the study period, I can go to the clinic where study staff are providing care, and that we will be examined and treated for free. I understand that if I become pregnant, I will withdraw the study.

I understand that I am free to take part in the study or refuse, and that I can withdraw myself from the study at any time, and without giving any reason. Deciding not to take part or to withdraw from the study will not affect the care that I or any of my family is normally entitled to.

I have had a chance to ask questions and have them answered.

Signature or thumb print of volunteer: _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

This form has been read by / I have read the above to _ _ _ _ _ _ _ _ _ _ _ _ _ _

(write name of volunteer)

in a language that she understands. I believe that she has understood what I explained and that she has freely agreed to take part in the study.

Signature of field worker: _

Name of field worker: _

Date: |__|__| / |__|__| / |__|__|__|__|

APPENDIX VI: Recipe collection form

FORM IG03 RECIPE COMPOSITION CALCULATION

Village _____ Subject's name: _____ WK ID: |__|__|__|__|__|__|__|__|__|__| Name of fieldworker _____

MDEG ID: |__|__|__| Today's date: |__|__|/|__|__|/|__|__| Sheet number/Total number (day) |__|/ |__| Recipe code: |__|-|__|__|-|__|__|__|__|

NAME OF THE RECIPE: _____

Weight of empty cooking saucepan (gr) |__|__|__|__| (@)

Name of ingredients added	Detailed description (Please note here if you have collected the labels for any of the ingredient)		Origin (*)	Weight (with or without container)	Weight of the container if used	Final weight of ingredient
Water added	1st addition (gr)	2nd addition (gr)	3rd addition (gr)	Total water:		

Total weight of all ingredients (#)_____

(*) (1)Home-garden, (2) bush, (3) shop, (4) rural market, (5) coast, (6) other –specify: |___| _____

Detailed instructions for the cooking

Cooking time-> Start: |__|__|:|__|__| Finish: |__|__|:|__|__|

Is saucepan cover? Yes/No

Is water boiling before introducing ingredients? Yes/No

Descriptive details of the cooking method

Finished recipe, ready to be eaten:

Total weight (full saucepan) (A)	Empty saucepan weight (@)	Final recipe weight (A)-(@)=(B)
|__|__|__|__|	|__|__|__|__|	|__|__|__|__|

How is this recipe consumed? Alone/With (staple/additional ingredients) _____

Water losses (gr):

Total weight of all ingredients (#)|__|__|__|__| - Final recipe weight (B) |__|__|__|__| = |__|__|__|__|

-When will the food be eaten? Date |__|__|/|__|__|/|__|__|

|__|__|/|__|__|/|__|__|

Meal: Breakfast/Lunch/Dinner

-Do you cook this recipe differently at different times of the year? (Explain what and how much)

-Number of people who will eat from this recipe: Children <5y |__|__| Adults |__|__|

-Receiving remittances? Yes/No

-Cooking place: closed kitchen/ ventilated kitchen/ open air

-Fuel type: wood/ dung/straw/ charcoal/ kerosene

Other comments.

GUIDANCE FOR USE

Subject will be visited the previous day to be reminded of our visit and to agree the arrival time of the observer.

Please request that she does not start cooking or eating anything before the observer arrives.

Observer will discuss with the subject whether her usual breakfast is re-heated food, or some other food (e.g. Bread with butter or porridge). If it is re-heated food prepared the previous day, the observer will visit her on the previous afternoon to attend the cooking of that recipe which will be used the following day.

If any of the foods (recipes or snacks) used is a processed product with a label on it (e.g. Chocolate bar, mayonnaise, sardines, etc.), please collect it and keep with the records of the subjects in her allocated folder.

RECIPES

1. Observer's name: Village assistant or Field worker and Supervisor if present
2. WK ID: West Kiang Number: provided in the list of subjects of every village, in the column on the left of the subjects name
3. MDEG ID: Study identification number |__|__|__|
   1st digit: Village of origin (1= Jiffarong; 2= Janneh Kunda; 3= Keneba)
   2nd & 3rd digit: Subject number (01-10 + additional recruitments)
4. Recipe code: The code of each recipe should be linked to the woman who prepared it
   |__|-|__|__|-|__|__|__|__|
   1st digit: Village of origin
   1= Jiffarong; 2= Janneh Kunda; 3= Keneba
   2nd – 4th digits: Subject number (01-10 + additional recruitments)
   5th & 7th digit: Recipe number (0001 to 9999)
5. Sheet no./Total no. sheets today|__|/ |__|
6. Before you weigh any food, make sure it is as ready for use (e.g. Vegetables have been peeled, fish has been skinned, unedible parts have been removed, etc.)
7. Description of foods: please, give as many details as possible (e.g. Type of fish/meat, with or without skin, etc.).
8. If the food is weighed directly on the scale, the cross the box for weight of container and write the same amount in the final weight column. If the food is weighed with a container, then weigh the container once it has been finished and subtract the weight of the container for the last column.
9. Use only the last 3 boxes of the final box for fried recipes (e.g. Pancakes, fishcakes, etc.)

APPENDIX VII: Dietary intake form

FORM IG04 24-H WEIGHED RECORDS: Village _____

Subject's name: _____ MDEG ID: |__|__|__|

Observer name: _____ WK ID: |__|__|__|__|__|__|__|__|__|__|

Today's date: |__|__|/|__|__|/|__|__| Sheet no./Total no. sheets today |__|/ |__|

Before eating: Re-heated? Yes/No Recipe code: |__|-|__|__|-|__|__|__|__|

Time weighed: |__|__|:|__|__|

Weight of container (gr) |__|__|__|__|

Container + 1st food (_____) |__|__|__|__|

Container + 1st + 2nd food (_____) |__|__|__|__|

Container + 1st + 2nd + 3rd food (_____) |__|__|__|__|

Container + 1st + 2nd + 3rd +4th food (_____) |__|__|__|__|

Container + 1st + 2nd + 3rd +4th + 5th food (_____) |__|__|__|__|

Comments on foods:

Meat/Fish flesh:

|__|__|__|__|

After eating: Time weighed: |__|__|:|__|__|

Weight of container+ food but without meat/fish bones:|__|__|__|__|

Specific left overs*: Meat/Fish bones weight |__|__|__|__|

Food:_____ |__|__|__|__|

Food:_____ |__|__|__|__|

Food:_____ |__|__|__|__|

Second attempt: Time weighed: |__|__|:|__|__|

Weight of container+ food but without meat/fish bones: |__|__|__|__|

Specific left overs*: Meat/Fish bones weight |__|__|__|__|

Food:_____ |__|__|__|__|

Food:_____ |__|__|__|__|

* Only if food items can be easily separated

Snacks

Snack 1 (Name/Weight): _____|__|__|__|__|

Snack 2(Name/Weight): _____|__|__|__|__|

Snack 3 (Name/Weight): _____|__|__|__|__|

Other comments: (subject sick, fasting, etc.)

DIETARY INTAKE

1. Observer's name: Village assistant or Field worker and Supervisor if present
2. WK ID: West Kiang Number: provided in the list of subjects of every village, in the column on the left of the subjects name
3. MDEG ID: Study identification number |___|___|___|

   1st digit: Village of origin

   A= Jiffarong; B= Janneh Kunda; C= Keneba

   2nd & 3rd digit: Subject number (01-10 + additional recruitments)
4. Sheet no./Total no. sheets today|___|/ |___|

   1st digit: Sheet number –to be filled 1-3 on every visit

   2nd digit: Total number of sheets- to be filled in at the end of the day.
5. Recipe code: |___|-|___|___|___|___|-|___|___|

   Copy the code from the recipe previously recorded.

   If no recipe was previously recorded, enter 0 in every box (it may be because the meal was prepared by someone else, or is re-heated from a previous day, or meal is made of bread with butter, etc.)

   It also should be noted to the supervisor when he/she collects the forms.
6. Food weight:

   Container + 1st + 2nd + 3rd food (_____)

   |___|___|___|___|

   Please, ask the subject to always add the meat/fish at the end

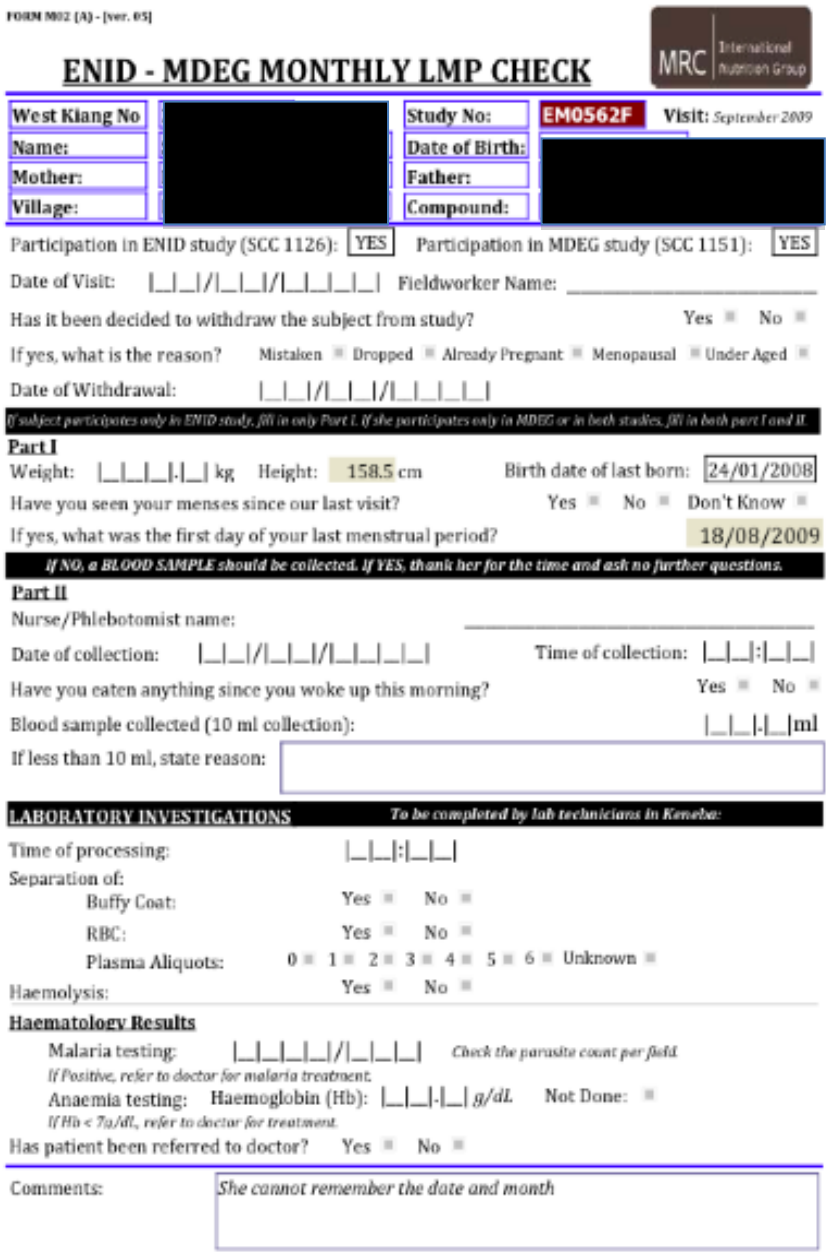
   On each line add in brackets the name of the new food added. If more space is needed, draw a star and continue in the "Comments on foods" section

   Fill the boxes with the weight in grams. Use the last boxes for it and fill in the empty ones with zero.

   Flesh only box should be filled in for the case of meat or fish and only once the left overs have been weighed. Then the weight of the bones should be subtracted from the total weight. If it is not applicable, leave empty.
7. "Specific left overs" section has to be filled in only for those particular ingredients which can be separated (e.g. Remaining fish, remaining bitter tomato) and zero (0000) should be written if nothing is left (e.g. Sauce finished)
8. Snacks: Subjects have to be followed the whole day to make sure every snack eaten in between meals is weighed and recorded. In this case, if the snack needs a container to be weighed, the scale should be put to zero with the empty container (tare system), before adding the snack.
9. Comments section: include here any information which can be useful to understand the data (e.g. Whether the women is fasting or sick, whether some adaptation had to be made to the procedure in order to cover for an unusual situation, etc.

NOTE: When writing the names of the foods, try to give as many details as possible (e.g. Name of fish, whether grounds are raw or roasted
If for any reason the snack weight coudln't be recorded because the orbserver was not present, then ask the subject and record in the form an estimation of the amoung (e.g. Big mango, handful of groundnuts, half tapalapa, etc.)

APPENDIX VIII: Menses questionnaire (blood collection after 1st missed menses)



FORM M02 (A) - [ver. 05]

MRC International nutrition Group

ENID - MDEG MONTHLY LMP CHECK

West Kiang No	[redacted]	Study No:	EM0562F	Visit: *September 2009*
Name:		Date of Birth:	[redacted]	
Mother:		Father:		
Village:		Compound:		

Participation in ENID study (SCC 1126): YES Participation in MDEG study (SCC 1151): YES

Date of Visit: |__|__|/|__|__|/|__|__|__|__| Fieldworker Name: _____

Has it been decided to withdraw the subject from study? Yes ▪ No ▪

If yes, what is the reason? Mistaken ▪ Dropped ▪ Already Pregnant ▪ Menopausal ▪ Under Aged ▪

Date of Withdrawal: |__|__|/|__|__|/|__|__|__|__|

If subject participates only in ENID study, fill in only Part I. If she participates only in MDEG or in both studies, fill in both part I and II.

Part I

Weight: |__|__|__|.|__| kg Height: 158.5 cm Birth date of last born: 24/01/2008

Have you seen your menses since our last visit? Yes ▪ No ▪ Don't Know ▪

If yes, what was the first day of your last menstrual period? 18/08/2009

If NO, a BLOOD SAMPLE should be collected. If YES, thank her for the time and ask no further questions.

Part II

Nurse/Phlebotomist name: _____

Date of collection: |__|__|/|__|__|/|__|__|__|__| Time of collection: |__|__|:|__|__|

Have you eaten anything since you woke up this morning? Yes ▪ No ▪

Blood sample collected (10 ml collection): |__|__|.|__|ml

If less than 10 ml, state reason:

LABORATORY INVESTIGATIONS *To be completed by lab technicians in Keneba:*

Time of processing: |__|__|:|__|__|

Separation of:

- Buffy Coat: Yes ▪ No ▪
- RBC: Yes ▪ No ▪
- Plasma Aliquots: 0 ▪ 1 ▪ 2 ▪ 3 ▪ 4 ▪ 5 ▪ 6 ▪ Unknown ▪

Haemolysis: Yes ▪ No ▪

Haematology Results

Malaria testing: |__|__|__|__|/|__|__|__| *Check the parasite count per field.*

If Positive, refer to doctor for malaria treatment.

Anaemia testing: Haemoglobin (Hb): |__|__|.|__| *g/dL* Not Done: ▪

If Hb < 7g/dL, refer to doctor for treatment.

Has patient been referred to doctor? Yes ▪ No ▪

Comments: *She cannot remember the date and month*

APPENDIX IX: Final version of the Indicator group paper accepted for publication at AJCN

Appendix X: Food composition tables

A. One-carbon metabolites and cofactors of common Gambian foods assessed (composition per 100g)

Ingredient	N	Folate (µg)[1]	Methionine (g)[2]	B12 (µg)[1]	B2 (mg)[2]	B6 (mg)[2]	Choline (mg)[3]	Betaine (mg)[3]
Atayah (green tea)	3	<5	<0.02	<0.20	<0.02	0.04	18.23	6.94
Aubergine	5	<5.0	0.05	<0.20	<0.01	<0.01	8.18	0.13
Baobab seeds	5	35	0.04	0.44	<0.05	0.04	13.36	0.89
Beans	5	82	0.36	0.39	0.03	0.18	155.80	3.83
Biscuits	5	<5.0	0.11	<0.20	<0.01	0.08	13.05	44.20
Bitter Tomato	5	11	0.03	0.24	0.02	0.05	18.03	0.04
Black pepper	4	<5.0	0.16	1.6	0.14	0.34	46.67	1.29
Bread	5	10	0.14	0.78	<0.01	0.03	12.89	44.32
Butternut squash	5	<5.0	<0.02	<0.20	<0.05	0.01	4.10	0.02
Cabbage	4	32	0.03	1.1	0.01	0.1	14.09	0.31
Cassava	5	7.2	0.03	5.9	<0.01	0.04	6.33	0.19
Chilli pepper, big	7	12	0.06	<0.20	0.14	0.4	16.63	3.23
Chilli pepper, dried	4	8.4	0.19	0.48	0.15	1.5	101.38	7.02
Coffee	5	42	0.12	1.9	<0.01	22.2	72.27	13.80
Condiments	5	<5.0	0.03	<0.20	<0.01	1.55	8.57	2.66
Egg, dry season	5	<5.0	0.3	1.2	0.22	0.03	231.57	0.83
Egg, wet season	5	13	0.48	2.9	0.36	0.05	214.30	1.06
Fish, challo, dried	4	<5.0	0.58	<0.20	0.15	0.14	50.36	27.35
Fish, challo, fresh	6	23	0.5	9.8	0.03	0.09	36.10	4.19
Fish, dried	6	9.7	0.54	21	0.08	0.11	43.84	22.93
Fish, jeja, lambasiso, tutuno and kotore, fresh	5	12	0.47	7.6	0.04	0.31	38.81	334.49
Fish, joto and furo, fresh	6	15	0.55	4.3	0.15	0.07	39.23	3.08
Fish, kujalo, kunkolongo and salango, fresh	5	15	0.65	3.4	0.07	0.07	48.95	43.04
Fish, moroko, kong, kungo and chalmoroo, fresh	5	23	0.56	17	0.07	0.17	43.19	22.26
Fish, pie	3	<5.0	0.13	1.4	<0.05	0.05	15.51	30.15

Ingredient	N	Folate (µg)[1]	Methionine (g)[2]	B12 (µg)[1]	B2 (mg)[2]	B6 (mg)[2]	Choline (mg)[3]	Betaine (mg)[3]
Fish, smoked	6	7.8	0.91	13	0.1	0.36	57.27	16.39
Fish, tambanjango, fresh	6	23	0.45	4.8	0.13	0.06	39.30	4.89
Garlic	3	6.8	0.06	0.37	<0.01	0.57	22.59	11.70
Groundnut, paste	6	21	0.32	0.24	<0.01	0.35	122.79	3.12
Groundnut, roasted	5	18	0.28	<0.20	<0.01	0.16	99.27	1.74
Groundnuts, cooked	7	11	0.25	<0.20	0.02	0.07	104.19	1.40
Groundnuts, raw, dry season	5	8.4	0.25	<0.20	<0.01	0.07	106.90	1.26
Groundnuts, raw, rainy season	5	14	0.28	<0.20	<0.01	0.08	119.05	1.89
Juice powder	3	<25	<0.02	1.5	<0.05	0.06	0.17	0.11
Kola nut	4	11	0.03	<0.20	<0.01	0.02	20.77	396.33
Leaves, baobab, dry (naa)	8	55	0.2	4.6	0.33	0.75	29.71	14.95
Leaves, cassava, fresh	5	35	0.17	0.58	0.22	0.17	32.54	1.75
Leaves, jambo nduroo, dried	3	26	0.51	1.7	0.73	1.41	157.10	6.93
Leaves, jambo nduroo, fresh	4	23	0.21	0.99	0.23	0.6	28.79	2.71
Leaves, kereng kereng, fresh	4	23	0.17	1.3	0.17	0.36	25.82	1.28
Leaves, kucha, dried	3	37	0.29	31	0.7	1.1	65.23	479.77
Leaves, kucha, fresh	8	8.1	0.09	0.3	0.06	0.2	13.81	45.87
Leaves, morongo, fresh	6	32	0.19	0.58	0.28	0.19	36.80	128.10
Leaves, onion, fresh	4	38	0.05	0.54	0.06	0.09	10.65	0.91
Leaves, sweet potato, fresh	5	26	0.17	1.3	0.15	0.13	20.34	16.95
Locus beans	4	210	0.15	3.9	0.19	1.6	35.41	2.32
Macaroni	5	<5.0	0.13	<0.20	<0.01	0.05	7.16	42.66
Maize, flour	5	13	0.13	<0.20	0.13	0.11	56.07	3.60
Maize, roast	5	27	0.15	<0.20	0.04	0.16	29.27	1.20
Maize, steamed	5	6.2	0.14	<0.20	<0.05	0.09	21.33	0.87
Mango, bush, dried	3	14	0.1	<0.20	0.07	0.06	25.00	18.78
Mango, fresh	5	14	0.03	<0.20	<0.05	0.06	7.87	0.16

Ingredient	N	Folate (µg)[1]	Methionine (g)[2]	B12 (µg)[1]	B2 (mg)[2]	B6 (mg)[2]	Choline (mg)[3]	Betaine (mg)[3]
Margarine	6	<5.0	<0.02	<0.20	<0.05	<0.01	2.30	0.00
Mayonnaise	5	<5.0	0.03	<0.20	0.06	<0.01	2.91	0.00
Meat beef, wet season	5	<5.0	0.33	1.5	0.04	0.12	44.34	7.86
Meat chicken	3	13	0.37	7	0.06	0.1	67.41	3.46
Meat, beef, dry season	3	21	0.4	34	0.14	0.12	88.36	6.42
Meat, goat	3	24	0.73	12	0.05	0.1	84.40	14.12
Milk, condensed	5	<5.0	0.15	0.48	0.35	0.05	44.80	1.82
Milk, fresh, dry season	5	7.2	0.1	0.6	0.22	0.06	14.12	1.52
Milk, fresh, wet season	2	11	0.19	0.39	0.51	0.07	31.47	2.63
Milk, powder	4	18	0.55	2.4	0.79	0.55	165.42	6.20
Milk, sour	5	<5.0	0.13	0.39	0.31	0.06	19.57	2.63
Millet, dry season	5	32	0.16	1.2	0.19	0.17	9.12	5.70
Millet, flour	5	24	0.13	0.47	0.06	0.16	20.06	3.29
Millet, porridge	5	15	0.1	0.9	0.04	0.07	10.89	1.84
Millet, wet season	5	13	0.17	0.49	0.05	0.17	14.22	2.58
Mustard	4	<5.0	0.12	<0.20	0.01	0.02	68.65	0.66
Nyet	1	13	0.84	12	0.17	0.27	37.67	5652.30
Okra	4	19	0.04	0.22	<0.01	0.02	11.11	4.90
Okra, powder	3	12	0.08	0.91	0.07	0.9	26.48	28.08
Onion	8	7.3	0.02	<0.20	<0.01	0.07	6.40	0.39
Palm oil	4	<5.0	<0.02	<0.20	<0.01	<0.01	0.62	0.78
Pancake	3	<5.0	0.1	0.38	<0.01	0.05	22.52	56.04
Papaya	5	8	0.02	<0.20	<0.05	0.12	6.76	0.12
Potato	4	12	0.06	0.5	<0.01	0.22	131.44	0.61
Pumpkin	5	6.6	0.02	0.41	<0.01	0.06	4.40	0.09
Rice, dempetengo (parched rice flakes)	3	13	0.18	0.97	<0.01	0.08	10.83	0.59
Rice, imported	5	10	0.21	0.89	<0.01	0.05	4.11	0.84
Rice, local	6	7.8	0.15	0.95	<0.01	0.04	3.09	0.14
Saf saf (condiment)	2	<5.0	<0.02	<0.20	<0.01	<0.01	0.00	0.00
Salt, lake	5	5.6	<0.02	<0.20	<0.01	<0.01	0.00	0.00
Sardines, canned	5	48	0.59	13	0.1	0.32	54.38	5.30
Seafood, crab, fresh	3	27	0.22	5.1	0.09	0.07	40.52	247.51

Ingredient	N	Folate (µg)[1]	Methionine (g)[2]	B12 (µg)[1]	B2 (mg)[2]	B6 (mg)[2]	Choline (mg)[3]	Betaine (mg)[3]
Seafood, oyster, fresh	3	13	0.33	32	0.03	<0.01	42.81	260.11
Seafood, shrimps, fresh	4	54	0.7	7.4	0.02	0.27	79.85	264.03
Seafood, shrimp, dry	3	9.9	0.71	1.4	<0.01	0.42	117.00	867.08
Sugar	4	<5.0	<0.02	0.24	<0.01	<0.01	0.00	0.58
Sweet Potato	5	7.3	0.04	0.93	<0.01	0.05	11.31	14.58
Tamarind	4	<5.0	0.04	<0.20	<0.01	0.11	7.51	0.46
Tea	5	<5.0	<0.02	<0.20	<0.02	0.02	34.30	4.44
Tomato	5	12	<0.02	<0.20	<0.01	0.16	11.37	0.16
Tomato paste	4	8.8	0.05	0.52	0.04	<0.60	44.82	0.64
Tomborungo (bush fruit)	5	<130	0.05	<0.20	<0.05	0.16	19.61	1.29
Vegetable oil	3	<5.0	<0.02	<0.20	<0.01	<0.01	0.00	0.00
Vinegar	4	<5	<0.02	<0.20	<0.01	<0.01	0.00	0.00
Wheat, flour	5	9.8	0.18	0.49	<0.01	0.04	11.71	81.22
Wonjo (hibiscus flower infusion)	3	<5	<0.02	<0.20	<0.02	<0.01	68.01	277.16

(1) Surface Plasmon Resonance Inhibition Assay

(2) HPLC

(3) LC/ESI IDMS

Note: Vitamin B12 assay is mainly designed to measure cyanocobalamin in vitamin B12-rich and fortified foods. Since vegetables are not main target matrices for the assay, it is suspected that vegetable products can give result measurable concentrations of vitamin B12 although not containing any vitamin. This is thought to be a matrix interference effect with the binding protein-surface interaction which usually occurs in samples which are blank for vitamin B12. Additionally, the extraction procedures may not be adequate for some of the vegetable samples tested and therefore, for the dietary intake analysis, foods of vegetal origin have been given value of 0 µg.

B. Total choline and choline compounds of common Gambian foods assessed (composition per 100g)

Ingredient	N	Total choline[a]	Free Choline	GPC	Pcho	PTC	SM
Atayah (green tea)	3	18.23	30.92	5.82	2.64	38.30	1.34
Aubergine	5	8.18	5.80	0.54	0.44	1.40	0
Baobab seeds	5	13.36	3.72	1.71	1.85	6.07	0
Beans	5	155.80	85.38	10.29	2.03	58.09	0
Biscuits	5	13.05	5.59	4.83	0	2.64	0
Bitter Tomato	5	18.03	12.73	1.01	0.36	3.93	0
Black pepper	4	46.67	19.02	1.58	0	1.98	0
Bread	5	12.89	6.67	5.09	0	1.12	0
Butternut squash	5	4.10	1.03	0.07	0.45	2.55	0
Cabbage	4	14.09	4.22	5.08	0.47	4.32	0
Cassava	5	6.33	1.42	0.05	1.20	3.66	0
Chilli pepper, big	7	16.63	14.55	0.86	0.34	0.88	0
Chilli pepper, dried	4	101.38	89.81	6.30	1.55	3.72	0
Coffee	5	72.27	76.75	3.34	0	0	0
Condiments	5	8.57	7.48	0.88	0	0.21	0
Egg, dry season	5	231.57	0.82	82.59	0.30	69.36	0
Egg, wet season	5	214.30	1.35	1.23	1.35	210.26	0.11
Fish, challo, dried	4	50.36	38.39	0.77	2.92	4.49	3.79
Fish, challo, fresh	6	36.10	5.20	5.56	2.31	20.03	3.00
Fish, dried	6	43.84	26.08	5.30	3.66	5.06	3.73
Fish, jeja, lambasiso, tutuno and kotore, fresh	5	38.81	4.58	4.46	1.30	19.24	7.75
Fish, joto and furo, fresh	6	39.23	2.58	3.32	2.25	27.19	3.89
Fish, kujalo, kunkolongo and salango, fresh	5	48.95	5.05	2.39	2.10	34.57	4.30

Ingredient	N	Total choline[a]	Free Choline	GPC	Pcho	PTC	SM
Fish, moroko, kong, kungo and chalmoroo, fresh	5	43.19	7.37	4.37	2.00	25.20	4.24
Fish, pie	3	15.51	10.09	2.50	1.43	1.49	0
Fish, smoked	6	57.27	9.67	5.89	3.15	34.79	3.77
Fish, tambanjango, fresh	6	39.30	1.74	7.64	1.28	24.94	3.69
Garlic	3	22.59	25.61	17.81	0.77	2.47	0
Groundnut, paste	6	122.79	70.26	2.88	3.10	46.55	0
Groundnut, roasted	5	99.27	54.07	2.03	4.40	38.76	0
Groundnuts, cooked	7	104.19	57.71	3.54	7.23	35.70	0
Groundnuts, raw, dry season	5	106.90	96.85	14.20	4.08	3.91	0
Groundnuts, raw, rainy season	5	119.05	72.22	26.01	4.55	4.10	0
Juice powder	3	0.17	0.09	0.08	0	0	0
Kola nut	4	20.77	13.32	1.47	1.79	3.48	0
Leaves, baobab, dry (naa)	8	29.71	27.89	1.39	1.09	1.61	0
Leaves, cassava, fresh	5	32.54	12.91	0.94	1.82	16.85	0
Leaves, jambo nduroo, dried	3	157.10	154.00	1.06	0.64	1.40	0
Leaves, jambo nduroo, fresh	4	28.79	17.74	1.21	0.39	9.44	0
Leaves, kereng kereng, fresh	4	25.82	1.90	0.40	0.96	22.54	0
Leaves, kucha, dried	3	65.23	53.37	21.31	102.71	6.48	0
Leaves, kucha, fresh	8	13.81	1.68	5.72	1.39	4.24	0

Ingredient	N	Total choline[a]	Free Choline	GPC	Pcho	PTC	SM
Leaves, morongo, fresh	6	36.80	14.85	0.64	1.17	16.42	0
Leaves, onion, fresh	4	10.65	5.07	0.48	0.88	4.22	0
Leaves, sweet potato, fresh	5	20.34	3.17	2.23	2.14	12.79	0
Locus beans	4	35.41	7.84	2.12	0.29	24.60	0.56
Macaroni	5	7.16	5.20	1.39	0	0.5	0
Maize, flour	5	56.07	55.21	0.26	0.50	0.09	0
Maize, roast	5	29.27	6.97	0.42	7.52	14.35	0
Maize, steamed	5	21.33	3.10	0.43	6.04	11.75	0
Mango, bush, dried	3	25.00	22.11	0.92	1.10	0.86	0
Mango, fresh	5	7.87	7.25	0.47	0	0.15	0
Margarine	6	2.30	0.09	1.95	0	0.26	0
Mayonnaise	5	2.91	0	0.96	0	1.95	0
Meat beef, wet season	5	44.34	0.97	10.33	0.47	28.25	4.05
Meat chicken	3	67.41	2.47	2.89	1.12	53.81	7.12
Meat, beef, dry season	3	88.36	4.58	19.00	1.22	56.97	6.58
Meat, goat	3	84.40	4.29	9.30	1.79	57.47	11.53
Milk, condensed	5	44.80	8.86	25.52	2.10	4.63	3.69
Milk, fresh, dry season	5	14.12	8.12	3.08	0	0.98	1.93
Milk, fresh, wet season	2	31.47	13.87	11.06	2.79	1.29	2.45
Milk, powder	4	165.42	27.12	108.54	9.62	12.11	8.02
Milk, sour	5	19.57	9.08	7.04	0	0.54	2.90
Millet, dry season	5	9.12	7.98	0.39	0.36	0.39	0
Millet, flour	5	20.06	17.81	2.14	0	0.10	0
Millet, porridge	5	10.89	10.32	0.29	0	0.28	0
Millet, wet season	5	14.22	12.37	1.39	0	0.46	0

Ingredient	N	Total choline[a]	Free Choline	GPC	Pcho	PTC	SM
Mustard	4	68.65	22.05	41.55	0.23	4.81	0
Nyet	1	37.67	29.32	0.37	6.33	4.43	0
Okra	4	11.11	1.48	0.81	0.64	8.18	0
Okra, powder	3	26.48	21.26	2.90	0.59	1.72	0
Onion	8	6.40	3.79	0.74	0.39	1.49	0
Palm oil	4	0.62	0.47	0.15	0	0	0
Pancake	3	22.52	7.65	13.59	0.24	1.03	0
Papaya	5	6.76	6.05	0.25	0.32	0.15	0
Potato	4	131.44	124.31	3.40	0.86	2.87	0
Pumpkin	5	4.40	1.02	0.14	0.26	2.97	0
Rice, dempetengo (parched rice flakes)	3	10.83	3.62	4.43	0.62	2.158	0
Rice, imported	5	4.11	2.17	1.71	0	0.23	0
Rice, local	6	3.09	0.80	1.09	0	1.20	0
Saf saf (condiment)	2	0.00	0	0	0	0	0
Salt, lake	5	0.00	0	0	0	0	0
Sardines, canned	5	54.38	2.48	5.00	1.20	41.31	4.38
Seafood, crab, fresh	3	40.52	3.67	12.01	1.10	19.64	4.57
Seafood, oyster, fresh	3	42.81	1.26	1.98	1.30	18.18	0.34
Seafood, shrimp, dry	4	79.85	17.09	25.10	4.15	50.95	13.81
Seafood, shrimps, fresh	3	117.00	4.95	2.69	2.46	60.02	9.73
Sugar	4	0.00	0	0	0	0	0
Sweet Potato	5	11.31	4.37	0.56	2.58	3.79	0
Tamarind	4	7.51	3.07	3.33	0.90	0.21	0
Tea	5	34.30	42.49	2.60	5.72	26.16	1.04
Tomato	5	11.37	9.114	0.71	0.60	0.95	0
Tomato paste	4	44.82	37.61	4.96	1.49	0.76	0

Ingredient	N	Total choline[a]	Free Choline	GPC	Pcho	PTC	SM
Tomborungo (bush fruit)	5	19.61	10.33	6.13	0.90	2.24	0
Vegetable oil	3	0.00	0	0	0	0	0
Vinegar	4	0.00	0	0	0	0	0
Wheat, flour	5	11.71	7.45	2.30	0	1.96	0
Wonjo (hibiscus flower infusion)	3	68.01	22.48	29.05	7.81	0.16	0

Method: LC/ESI IDMS

[a]Calculated as sum of the 5 choline compounds

Units: mg

APENDIX XI: Food collection forms

FORM IG02A –FOOD COLLECTION INFORMATION SHEET

B) Identification of Food: information to be collected

- Sample code number: |__|__|__|__|__|
- Common name of food (+alternative names, scientific name): _____
  _____
- Part of plant/animal ☐ (which part_____); Entire ☐
- State of maturity: immature ☐; ripe ☐
- Storage conditions when collected: _____
- Labels and ingredients list: _____
  _____
  Process and preservation method: canned ☐ smoked ☐ sun-dried ☐
  other ☐ _____
- Other comments: (e.g. Unusual food- Take picture!) _____
  _____

C) Record of collection:

- Date of collection: |__|__|/|__|__|/|__|__|
- Time of collection: |__|__|:|__|__|
- Name of collector: _____
- Sampling point type: field ☐, garden ☐, rural market ☐, shop ☐, other ☐ specify:_____
- Place of origin (village): Jiffarong ☐, Janneh Kunda ☐ Keneba ☐
  Other ☐ specify: _____
- Weight of food collected: |__|__|__|__| (gr)
- Number of items: |__|__|
- Other relevant details: _____
- Preparation method for consumption (cooking method, time of cooking, addition of water/oil, etc):
  _____
  _____

FORM IG02B –LABORATORY FOOD PROCESSING FORM

- Sample code number: |__|__|__|__|__|
- Processing date: |__|__|/|__|__|/|__|__|
- Name of person handling sample: _____

Cooking

- Part of the food processed (preparation/cleaning): _____

_____

- Cooking Yes ☐ No (raw) ☐
- If boiled: Time |__|__|__|
- Weight before cooking: |__|__|__| (gr) + added water |__|__|__| (gr)
- Weight after cooking: |__|__|__| (gr)

Processing

- Weight of blended ingredient for sample: |__|__|__| (gr)
- Weight of blended ingredient and container: |__|__|__| (gr)
- Date/time of freeze-drying start: |__|__|/|__|__|/|__|__| |__|__|:|__|__|
  - Weight at 0 h: |__|__|__| (gr)
  - Weight at 12 h: |__|__|__| (gr)
  - Weight at 24 h: |__|__|__| (gr)
  - Weight at 36 h: |__|__|__| (gr)
  - Weight at 48 h: |__|__|__| (gr)
  - Other: Hours |__|__|
- Date/time of freeze-drying end: |__|__|/|__|__|/|__|__| |__|__|:|__|__|

Storage

- At -40ºC ☐
- Comments:

_____

_____

Preparation of composite sample (check Excel template for calculation on % and reconstitution)

- Mixing date: |__|__|/|__|__|/|__|__|
- Sample code of composite sample: |__|__|__|__|__|

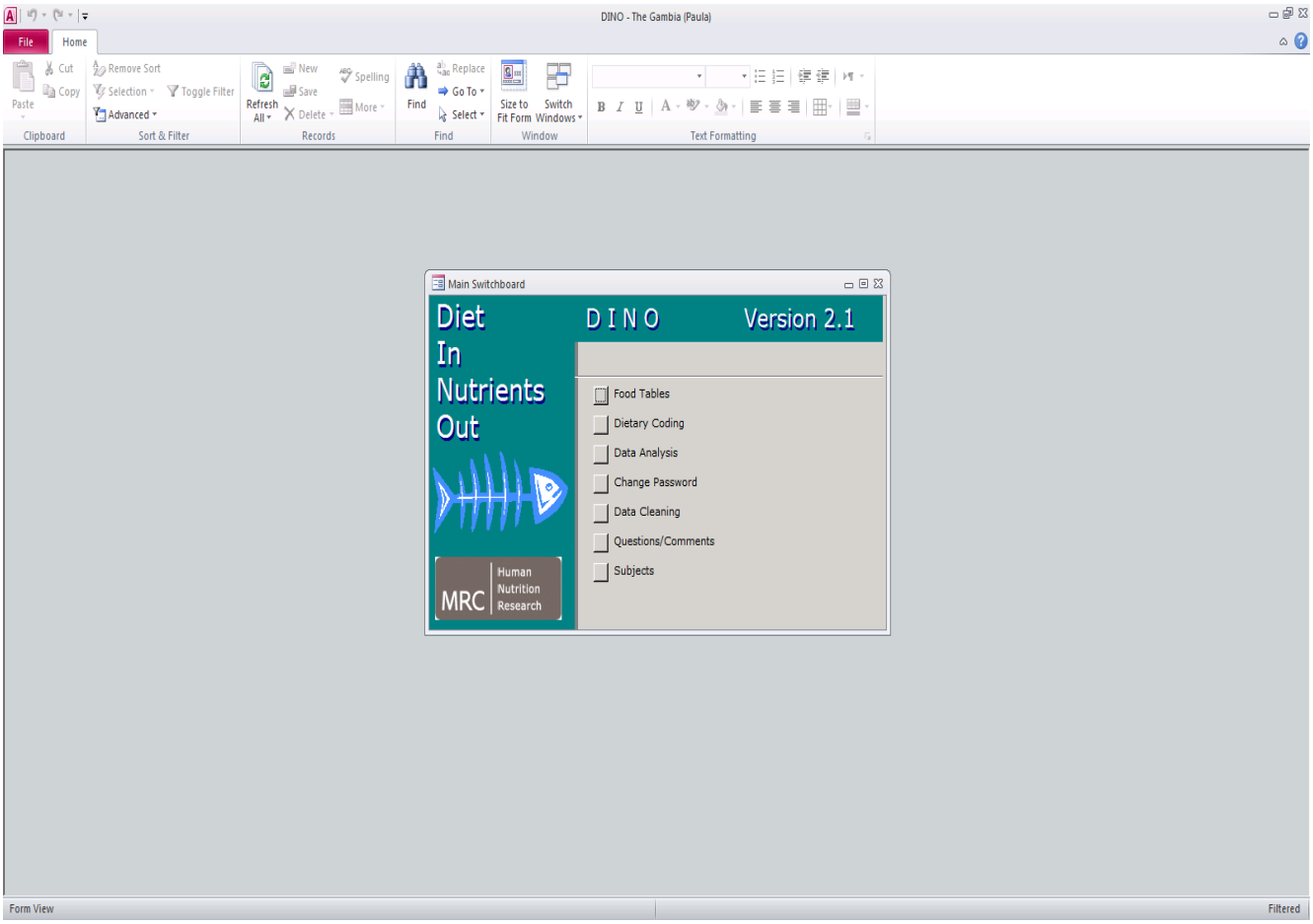
Shipping

- Date of dispatch for overseas laboratory: |__|__|/|__|__|/|__|__|
- Need for water reconstitution on arrival: |__|__|__| (gr)
- Comments: _____

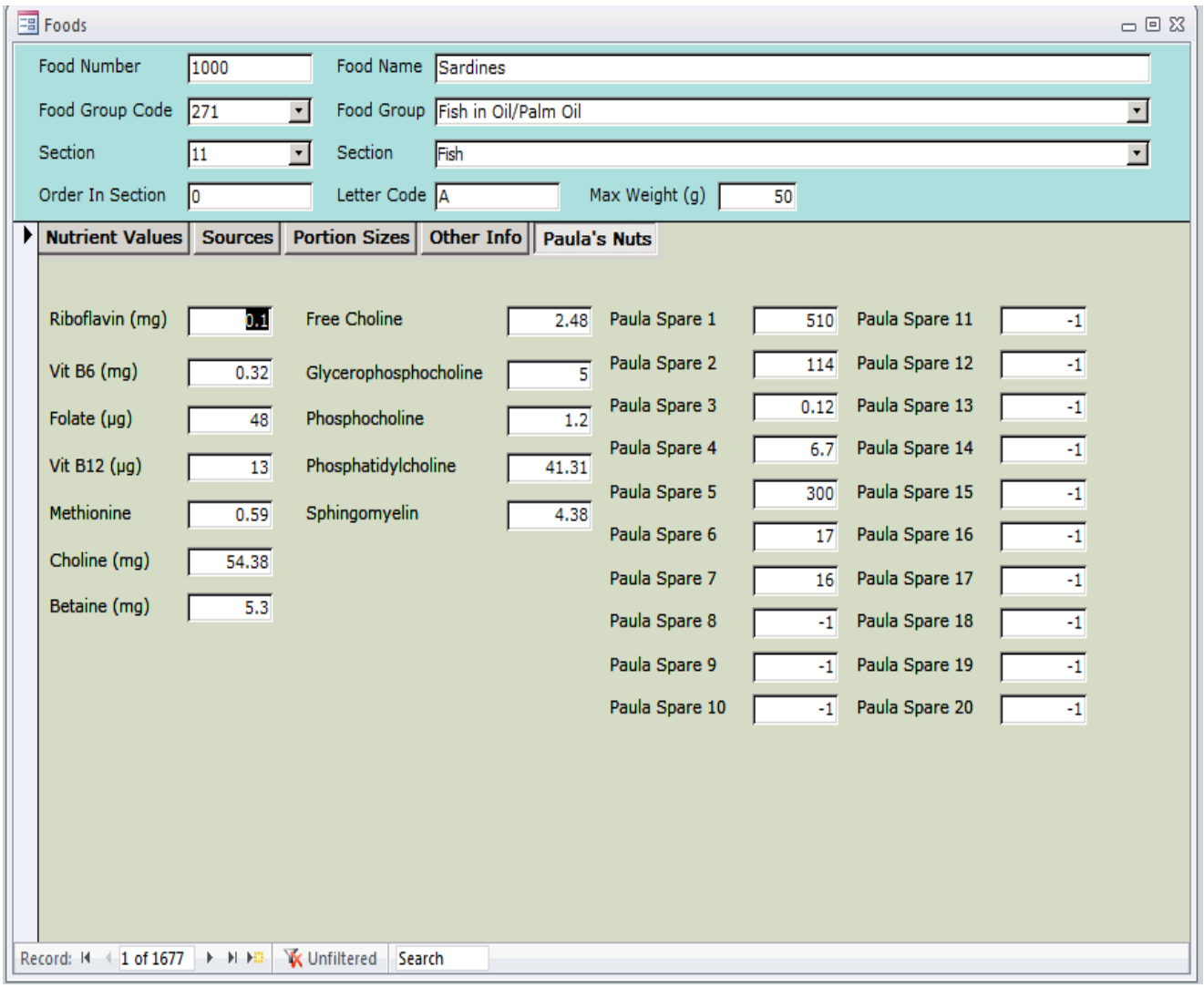
_____

APPENDIX XII: Examples of DINO data entry screens

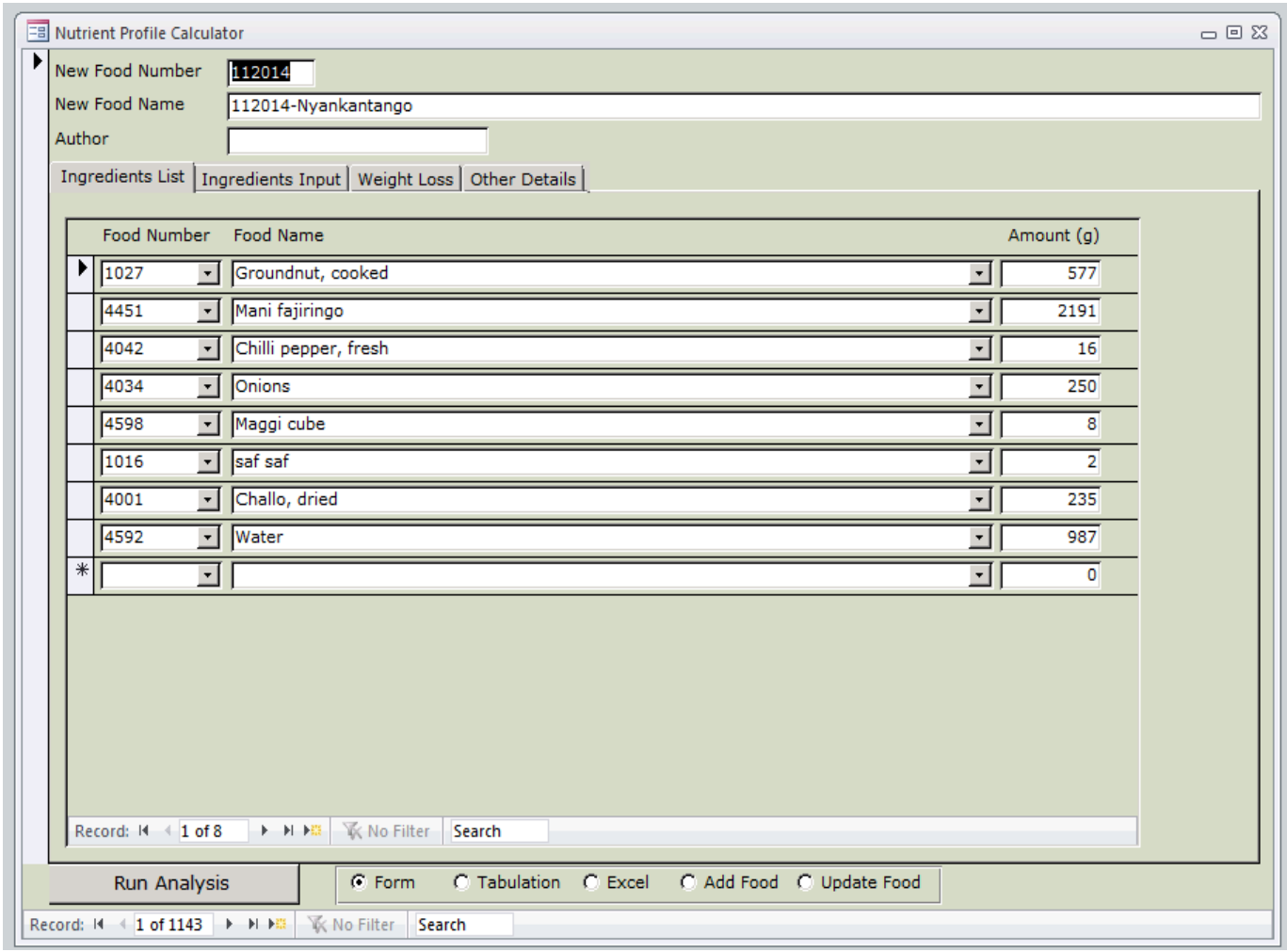
a) Main menu



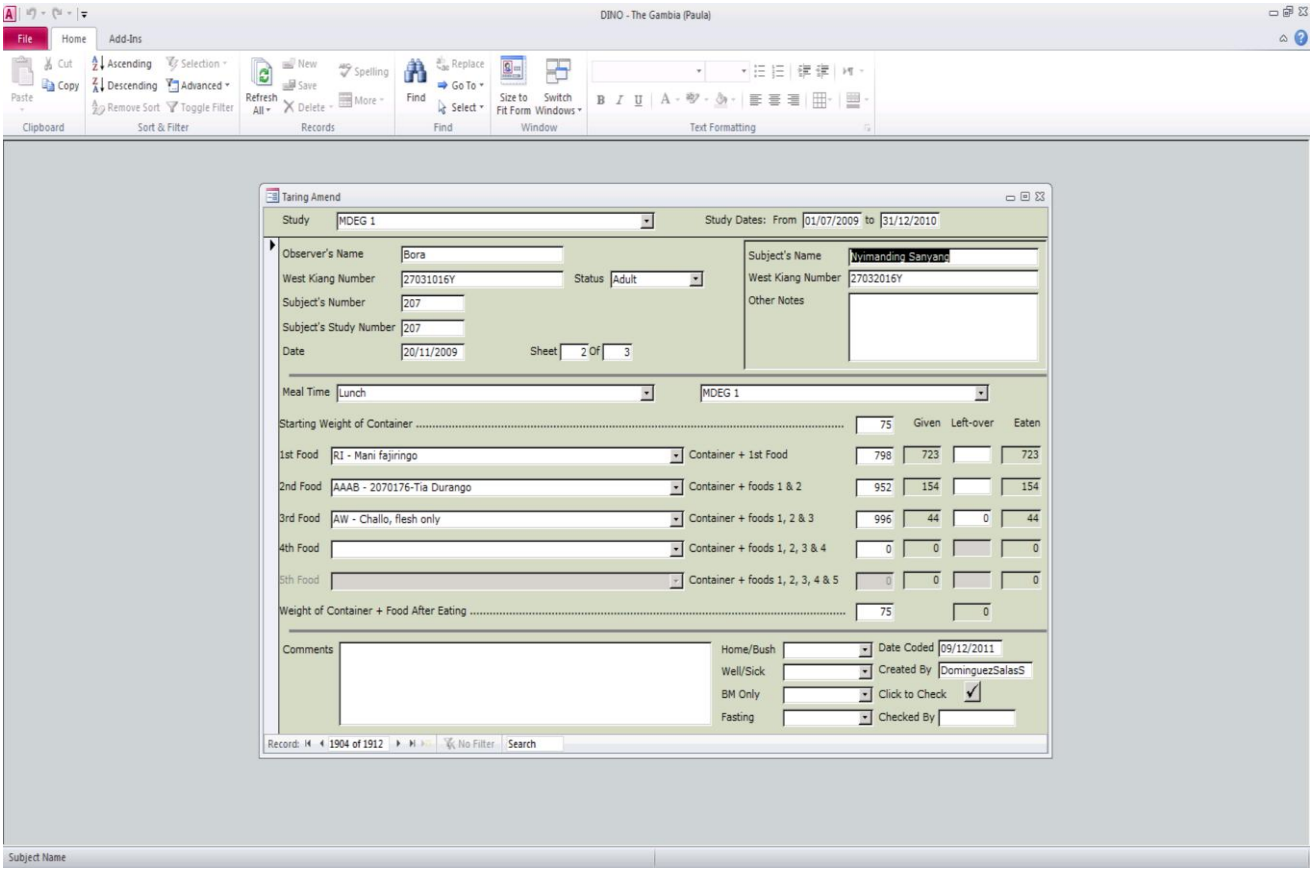
b) Food composition



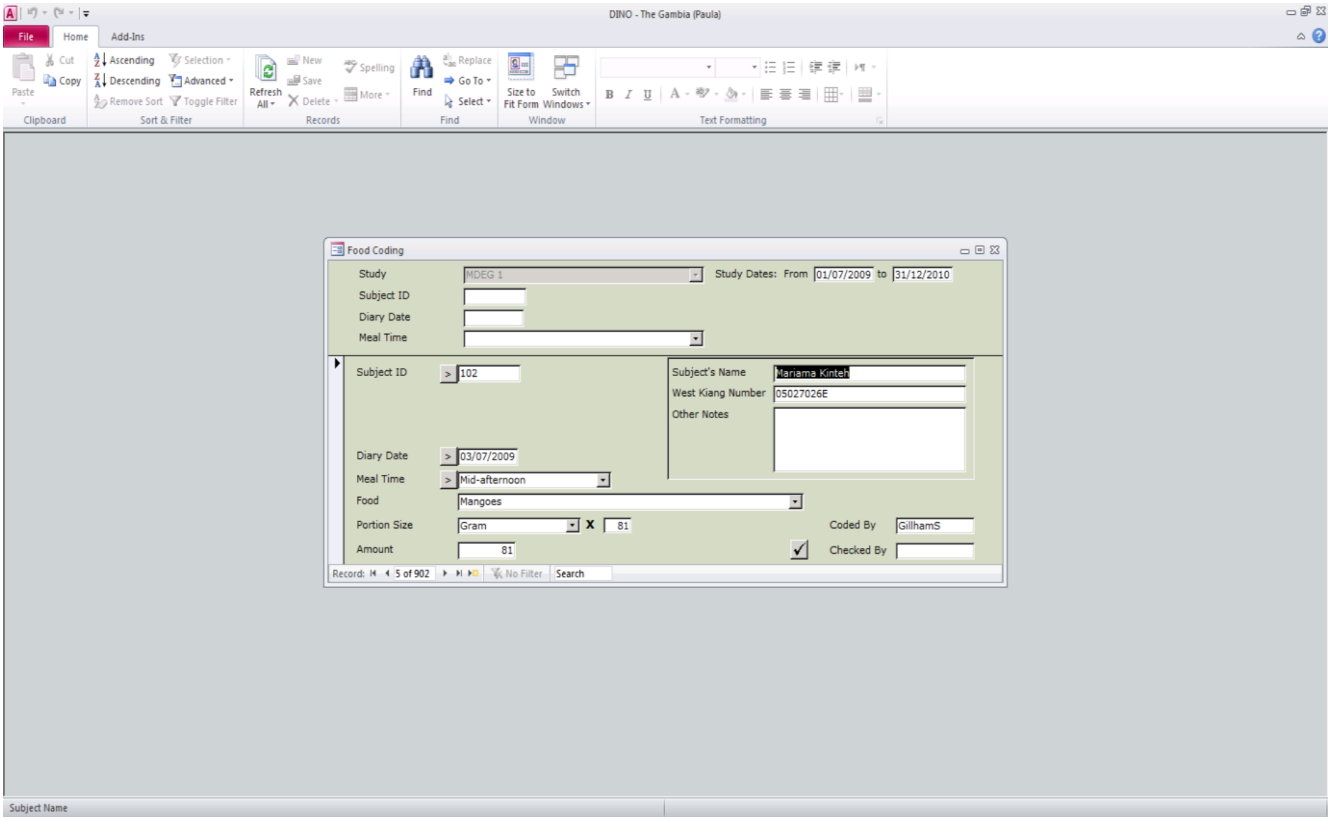
c) Recipe


Nutrient Profile Calculator
New Food Number
112014
New Food Name
112014-Nyankantango
Author
Ingredients List
Ingredients Input
Weight Loss
Other Details
Food Number
Food Name
Amount (g)
1027
Groundnut, cooked
577
4451
Mani fajiringo
2191
4042
Chilli pepper, fresh
16
4034
Onions
250
4598
Maggi cube
8
1016
saf saf
2
4001
Challo, dried
235
4592
Water
987
0
Record: 1 of 8
No Filter
Search
Run Analysis
Form
Tabulation
Excel
Add Food
Update Food
Record: 1 of 1143
No Filter
Search

d) Dietary intake



e) Snacks


DINO - The Gambia (Paula)
File
Home
Add-Ins
Food Coding
Study
MDEG 1
Study Dates: From 01/07/2009 to 31/12/2010
Subject ID
Diary Date
Meal Time
Subject ID 102
Subject's Name Mariama Kinteh
West Kiang Number 05027026E
Other Notes
Diary Date 03/07/2009
Meal Time Mid-afternoon
Food Mangoes
Portion Size Gram X 81
Coded By GillhamS
Amount 81
Checked By
Record: 5 of 902
No Filter
Search
Subject Name

APPENDIX XIII: Menses questionnaire (urine collection after 2nd missed menses)

FORM M02 (B) - [Ver. 02]

ENID - MDEG MONTHLY LMP CHECK AFTER FIRST MENSES LOSS (SECOND VISIT)

MRC International Nutrition Group

West Kiang No:		Study ID:		Visit: #Error
Name:		Date of Birth:		
Mother:		Father:		
Village:		Compound:		

Date of Visit: |__|__|/|__|__|/|__|__|__|__|

Fieldworker Name: _____

Has it been decided to withdraw the subject from study? Yes ☐ No ☐

If yes, what is the reason?

☐ Self Withdrawal ☐ Moved Away ☐ Reaches Menopause ☐ Died ☐ Other: _____

Date of Withdrawal: | | |/| | |/| | | | |

Weight: |__|__|__|.|_ kg

Height: cm

Have you seen your menses in the last month? Yes ☐ No ☐ Don't Know ☐

If yes, what is the first day of your last menstrual period? |__|__|/|__|__|/|__|__|__|__|

If no, fieldworker to decide whether urine sample should be collected? Yes ☐ No ☐

If NO, (Second Loss), a URINE SAMPLE should be collected for dipstick pregnancy testing.

Urine sample collected? Yes ☐ No ☐

Date of Collection: |__|__|/|__|__|/|__|__|__|__|

Time of Collection: |__|__|:|__|__|

LABORATORY RESULTS *To be completed by the lab technicians in Keneba:*

Dipstick Results: Positive ☐ Negative ☐ Repeat Sample ☐

Date of Processing: __|__|/|__|__|/|__|__|__|__|

Time of Processing: |__|__|:|__|__|

Comments: