INFLUENCE OF B-VITAMINS ON ONE-CARBON METABOLISM AND ASSOCIATIONS WITH CANCER RISK AND REPRODUCTIVE STATE

A Dissertation
Presented to the Faculty of the Graduate School of Cornell University
In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

by
Sajin Bae
May 2017
INFLUENCE OF B-VITAMINS ON ONE-CARBON METABOLISM AND ASSOCIATIONS WITH CANCER RISK AND REPRODUCTIVE STATE

Sajin Bae, Ph.D.
Cornell University 2017

Folate, choline and vitamin B12 are essential nutrients involved in one-carbon metabolism (OCM), a network of interconnected pathways necessary for the de novo synthesis of purines and thymidylate and for the remethylation of homocysteine to methionine. Disruptions in OCM are associated with aberrant DNA synthesis and methylation and high risk for cancer. Thus, it is of particular importance to elucidate the role of these nutrients in the functioning of OCM. In addition, the status of these nutrients and their demand differ by reproductive and/or pathological state, further addressing the need to better understand the effects of these nutrients.

This dissertation research involves both human participant studies and laboratory-based molecular research to advance current knowledge of the role of these nutrients.

Study 1 examined the impact of mandatory folic acid fortification on DNA methylation status among postmenopausal women enrolled in the Women’s Health Initiative Observational Study (WHI-OS). As expected, given the role of folate in OCM, women with higher red blood cell (RBC) folate concentration had higher DNA methylation in the pre-fortification period. However, this expected result was not observed in the post-fortification period during which women with higher (vs. lower) RBC folate status had lower DNA methylation. Overall, these findings suggest an inverted U-shaped relationship between folate status and DNA methylation across fortification periods, and further investigation is warranted to clarify the health outcomes of the inverse relationship observed in the era of folic acid fortification.

Study 2 examined associations between biomarkers of choline metabolism and colorectal
cancer risk in a case-control study nested within the WHI-OS. The main findings indicate 1) a positive association between plasma trimethylamine N-oxide (TMAO; a derivative of choline produced by intestinal bacteria) and rectal cancer risk; and 2) an inverse relationship between plasma betaine and colorectal cancer risk. These findings demonstrate that alterations in choline metabolism associate with higher risk of colorectal cancer, suggesting the potential utilization of these biomarkers as predictors of increased colorectal cancer risk.

Study 3 assessed changes in status and functional biomarkers of vitamin B12 among pregnant, lactating and control (nonpregnant, nonlactating) women who consumed equivalent vitamin B12 intakes under controlled feeding conditions. Pregnant (vs. control) women had a higher ratio of plasma holotranscobalamin (bioactive form of vitamin B12) to total vitamin B12, indicating that greater amounts of vitamin B12 are partitioned toward the biologically active form in this reproductive state. Overall, the results of this study suggest that metabolic adaptations transpire to enhance vitamin B12 supply during pregnancy.

Study 4 employed in vitro cell culture models to investigate the effect of folate-independent generation of formate, a primary source of one-carbons for folate-mediated OCM, on the synthesis of purines and thymidylate. The study findings demonstrate that in human hepatocarcinoma (HepG2) cells, alcohol dehydrogenase 5 is a source of formate for de novo purine biosynthesis, especially during folate deficiency when folate-dependent formate production is limited.

Taken together, this dissertation research spans from in vitro molecular studies to epidemiological studies to address the role of folate, choline and vitamin B12. The findings of this research will help inform the development of nutrient intake recommendations and the use of nutritional biomarkers for disease prediction.
BIOGRAPHICAL SKETCH

Sajin Bae grew up in Busan, South Korea. She earned a B.S. degree in the Department of Biological Resources and Technology at Yonsei University, Wonju, South Korea in 2010. After finishing her B.S. degree, she entered the M.S. degree program in the Department of Animal Sciences at North Dakota State University. She worked as a graduate research assistant in Dr. Chung Park’s laboratory, and her thesis work focused on investigating the effect of maternal methyl diet on epigenetic modification and mammary cancer risk by using a rodent model. After finishing her M.S. degree in 2012, she entered the Ph.D. degree program in Nutritional Sciences at Cornell University to advance her training in the field of biomedical nutrition research. Sajin joined the laboratory of Dr. Marie Caudill in the spring of 2012 and completed three projects investigating the role of folate, choline and vitamin B12 in one-carbon metabolic pathways and their impacts on cancer risk in humans. In 2014, she was selected as a trainee on the National Institutes of Health (NIH)-sponsored T32 Predoctoral Training Program in Nutrition. Through this program, she extended her research into laboratory-based molecular studies under the direction of Dr. Patrick Stover to investigate origins of endogenous formate and their contribution to one-carbon metabolism by using in vitro cell culture models. Building on her multidisciplinary research experience in both human participant studies and laboratory-based molecular research, she participated in the 2015 WHO/Cochrane/Cornell Summer Institute for Systematic Reviews in Nutrition for Global Policy Making program and is currently taking the lead on developing a Cochrane systematic review. In addition to her dedication as a researcher, Sajin has served as a teaching assistant for several undergraduate nutrition courses and was awarded the Outstanding Graduate Teaching Assistant Award from the College of Agriculture and Life Sciences, Cornell University.
Dedicated to my Father, husband, daughter and parents
ACKNOWLEDGEMENTS

First and foremost, I would like to thank God for His sincere love that sustains me in times of joy and trouble. I thank Him for being my guide in each and every step I have taken and will take in my life. I would like to express my deepest love to my husband Kevin Lee who always stands by me and supports me with prayers, love and encouragement. Special thanks to my daughter Evelyn, the best blessing given to me and Kevin. I would like to thank my parents, sister and brother, whom I love with all my heart, for believing in me and being there with me whenever I need. To all my friends, thank you for your support and sharing at every moment of this journey.

I would like to thank Dr. Marie Caudill and Dr. Patrick Stover for their constant guidance, support and advice and for setting a high standard of excellence in research, teaching and mentoring. I have learned so many things from you, which are integral in my career path. I also thank my committee members, Dr. Rebecca Seguin and Dr. Jeffery Sobal for their supervisory roles and kind support for my research and academic accomplishments. Thanks to the members of the Caudill lab and Stover lab for their sharing and support, and I truly enjoyed working with everyone in the lab. Thanks to all who have given me sincere care and support in the Division of Nutritional Sciences at Cornell University.

Research reported in this publication was in part supported by the National Institutes of Health under award T32-DK007158. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) or the National Institutes of Health.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOGRAPHICAL SKETCH</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>PREFACE</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER 1: Impact of folic acid fortification on global DNA methylation and one-carbon biomarkers in the Women’s Health Initiative Observational Study cohort</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2: Plasma Choline Metabolites and Colorectal Cancer Risk in the Women’s Health Initiative Observational Study</td>
<td>33</td>
</tr>
<tr>
<td>CHAPTER 3: Vitamin B-12 status differs among pregnant, lactating, and control women with equivalent nutrient intakes</td>
<td>70</td>
</tr>
<tr>
<td>CHAPTER 4: Alcohol dehydrogenase 5 is a source of formate for de novo purine biosynthesis in HepG2 Cells</td>
<td>104</td>
</tr>
<tr>
<td>AFTERWORD</td>
<td>133</td>
</tr>
<tr>
<td>APPENDIX B: Copy right approval forms</td>
<td>191</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

CHAPTER 1

SUPPLEMENTARY FIGURE S1.1 ................................................................. 27

CHAPTER 3

FIGURE 3.1 .......................................................................................... 87
FIGURE 3.2 .......................................................................................... 90

CHAPTER 4

FIGURE 4.1 .......................................................................................... 108
FIGURE 4.2 .......................................................................................... 116
FIGURE 4.3 .......................................................................................... 117
FIGURE 4.4 .......................................................................................... 119
SUPPLEMENTAL FIGURE S4.1 ............................................................... 130
SUPPLEMENTAL FIGURE S4.2 ............................................................... 131
SUPPLEMENTAL FIGURE S4.3 ............................................................... 132
LIST OF TABLES

CHAPTER 1

TABLE 1.1..............................................................................................................6
TABLE 1.2.............................................................................................................8
TABLE 1.3...........................................................................................................9
TABLE 1.4..........................................................................................................10
TABLE 1.5.........................................................................................................11
SUPPLEMENTARY TABLE S1.1.................................................................28
SUPPLEMENTARY TABLE S1.2.................................................................29
SUPPLEMENTARY TABLE S1.3.................................................................30
SUPPLEMENTARY TABLE S1.4.................................................................31
SUPPLEMENTARY TABLE S1.5.................................................................32

CHAPTER 2

TABLE 2.1........................................................................................................43
TABLE 2.2........................................................................................................46
TABLE 2.3........................................................................................................47
TABLE 2.4........................................................................................................49
TABLE 2.5.......................................................................................................51
TABLE 2.6.......................................................................................................53
TABLE 2.7.......................................................................................................54
SUPPLEMENTARY TABLE S2.1.............................................................67
SUPPLEMENTARY TABLE S2.2.............................................................68
SUPPLEMENTARY TABLE S2.3.............................................................69

x
CHAPTER 3

TABLE 3.1...............................................................................................................82
SUPPLEMENTAL TABLE S3.1................................................................................101

CHAPTER 4

TABLE 4.1...............................................................................................................121
PREFACE

One-carbon metabolism is a metabolic network essential for the *de novo* synthesis of purines and thymidylate and for the remethylation of homocysteine to methionine. Disruptions in one-carbon metabolism, which can arise from insufficient intake of relevant nutrients (including folate, choline and vitamin B12) and/or genetic variants, have been linked to a number of human diseases including cancer, neurodegenerative diseases and developmental anomalies. The status of these nutrients and their demand are dependent on each other and can be altered by reproductive and/or pathological state. The overarching goal of this dissertation research was to investigate the associations of the intake and status of these nutrients with 1) the functioning of one-carbon metabolic pathways; 2) cancer risk; and 3) reproductive state. To achieve this goal, a multidisciplinary translational approach including both human participant studies and *in vitro* cell culture experiments was used with a focus on the following specific aims:

**Aim 1**: To investigate the impact of mandatory folic acid fortification on DNA methylation status and the relationship between DNA methylation and one-carbon metabolic biomarkers. This aim was achieved by using samples from postmenopausal women of the Women’s Health Initiative Observational Study (WHI-OS) cohort collected before and after folic acid fortification. Study findings are presented in chapter 1.

**Aim 2**: To investigate associations between plasma biomarkers of choline metabolism and colorectal cancer risk among postmenopausal women in a case-control study nested within the WHI-OS. Results are presented in chapter 2.

**Aim 3**: To investigate changes in vitamin B12 status during pregnancy and lactation under controlled feeding conditions. This aim was achieved by assessing and comparing vitamin B12 status and functional biomarkers among pregnant, lactating and control (nonpregnant,
nonlactating) women who consumed equivalent vitamin B12 intakes under controlled feeding conditions. Study findings are presented in chapter 3.

**Aim 4:** To investigate the contribution of folate-independent generation of formate, a primary source of one-carbons for folate-mediated one-carbon metabolism, to the *de novo* synthesis of purines and thymidylate. This aim was achieved by using *in vitro* cell culture models that employed a gene knockout system and radioactive isotope tracers. Results are presented in chapter 4.

**Aim 5:** To evaluate and summarize scientific evidence on the effects of provision of folic acid for reducing arsenic toxicity in arsenic-exposed children and adults. This is an ongoing systematic review being done with the Cochrane Developmental, Psychosocial and Learning Problems Group. The results of this review will serve to inform evidence-based public health guidelines on folic acid interventions in arsenic-exposed populations including all ages and gender groups. The Cochrane protocol, which prespecifies a detailed plan for the review and is required to be published prior to conducting a Cochrane review, has been developed and is presented in Appendix A.

This dissertation research yielded four published primary research manuscripts in peer-reviewed journals (Chapters 1-4) and one Cochrane Systematic Protocol, which has been published in the *Cochrane Database of Systematic Reviews* (Appendix A).
CHAPTER 1

Impact of folic acid fortification on global DNA methylation and one-carbon biomarkers in the Women’s Health Initiative Observational Study cohort*

ABSTRACT

DNA methylation is an epigenetic mechanism that regulates gene expression and can be modified by one-carbon nutrients. The objective of this study was to investigate the impact of folic acid (FA) fortification of the US food supply on leukocyte global DNA methylation and the relationship between DNA methylation, red blood cell (RBC) folate, and other one-carbon biomarkers among postmenopausal women enrolled in the Women’s Health Initiative Observational Study. We selected 408 women from the highest and lowest tertiles of RBC folate distribution matching on age and timing of the baseline blood draw, which spanned the pre-(1994–1995), peri- (1996–1997), or post-fortification (1998) periods. Global DNA methylation was assessed by liquid chromatography-tandem mass spectrometry and expressed as a percentage of total cytosine. We observed an interaction ($P = 0.02$) between fortification period and RBC folate in relation to DNA methylation. Women with higher (vs. lower) RBC folate had higher mean DNA methylation (5.12 vs. 4.99%; $P = 0.05$) in the pre-fortification period, but lower (4.95 vs. 5.16%; $P = 0.03$) DNA methylation in the post-fortification period. We also observed significant correlations between one-carbon biomarkers and DNA methylation in the pre-fortification period, but not in the peri- or post-fortification period. The correlation between plasma homocysteine and DNA methylation was reversed from an inverse relationship during the pre-fortification period to a positive relationship during the post-fortification period. Our data suggest that (1) during FA fortification, higher RBC folate status is associated with a reduction in leukocyte global DNA methylation among postmenopausal women and; (2) the relationship between one-carbon biomarkers and global DNA methylation is dependent on folate availability.
INTRODUCTION

DNA methylation is an epigenetic modification of the genome, which influences gene expression and genome integrity.\(^1\) DNA methylation can be modified by nutrients involved in one-carbon metabolism (e.g., folate, choline, vitamin B12, and vitamin B6), and disturbances in methylation reactions caused by abnormal status of these nutrients have been implicated in a number of human diseases including cancer.\(^2\)–\(^6\) Folate, in the form of 5-methyltetrahydrofolate (5-methyl-THF), participates in cellular methylation reactions (including DNA methylation) by donating a methyl group for the vitamin B12-dependent re-methylation of homocysteine to methionine (Supplementary Figure S1.1). Folic acid (FA), the synthetic form of folate, can also participate in DNA methylation after its reduction to THF and conversion to 5-methyl-THF.\(^7\) Homocysteine re-methylation to methionine can also proceed via a folate and B12-independent route in which betaine (a derivative of choline) serves as the methyl donor.\(^8\)

In January of 1998, the US Food and Drug Administration mandated FA fortification of enriched cereal-grain products (i.e., addition of 140 μg of FA/100 g of grain) in an effort to reduce the occurrence of neural tube defects, a mandate that some food companies initiated in 1996 and 1997.\(^9\) FA fortification of the US food supply has led to significant increases in serum and red blood cell (RBC) folate concentrations as well as decreases in plasma total homocysteine;\(^10,11\) however, less is known about the impact of FA fortification on DNA methylation.

In this report, we investigated the association between mandatory FA fortification and leukocyte global DNA methylation, as well as the relationship between global DNA methylation, RBC folate, plasma choline, and other biomarkers of one-carbon metabolism among postmenopausal women enrolled in the Women’s Health Initiative Observational Study (WHI-
OS).
RESULTS

Characteristics of the study population

The participants of this study were a subset of the control group from a nested case-control study investigating colorectal cancer risk in the WHI-OS.\textsuperscript{12,13} Baseline demographic and biochemical characteristics of the participants, which corresponded to FA fortification periods [pre (1994–1995), peri (1996–1997), or post (1998)], are shown in Table 1.1. BMI differed among FA fortification periods ($P = 0.01$) with a higher mean BMI (28.3 kg/m\(^2\)) in the post-fortification period than in the pre- (26.5 kg/m\(^2\); $P = 0.02$) and peri-fortification (26.4 kg/m\(^2\); $P = 0.01$) periods. The ethnic distribution also differed among fortification periods ($P = 0.03$).

Plasma folate differed among FA fortification periods ($P = 0.002$) with higher median plasma folate (20 ng/mL) in the post-fortification period than in the pre- (14 ng/mL; $P = 0.001$) and peri-fortification (16 ng/mL; $P = 0.002$) periods. Similarly, RBC folate differed among FA fortification periods ($P = 0.002$) with higher median RBC folate in the peri- (572 ng/mL; $P = 0.02$) and post-fortification (726 ng/mL; $P < 0.001$) periods compared with the pre-fortification (424 ng/mL) period.
Table 1.1. Baseline characteristics of the study participants (n = 408) according to folic acid (FA) fortification period\(^1\)\(^-\)\(^3\)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)(^4)</td>
<td>67 ± 7</td>
<td>67 ± 7</td>
<td>67 ± 6</td>
<td>0.87</td>
</tr>
<tr>
<td>BMI (kg/m(^2))(^4)</td>
<td>26.5 ± 5(^a)</td>
<td>26.4 ± 5(^a)</td>
<td>28.3 ± 6(^b)</td>
<td>0.01</td>
</tr>
<tr>
<td>Race/ethnicity(^5)</td>
<td></td>
<td></td>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>86</td>
<td>88</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>14</td>
<td>12</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Education(^5)</td>
<td></td>
<td></td>
<td></td>
<td>0.69</td>
</tr>
<tr>
<td>≤ High school</td>
<td>21</td>
<td>24</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>≥ College</td>
<td>79</td>
<td>76</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Pack-years of smoking(^6)</td>
<td>0 (0–10)</td>
<td>0 (0–15)</td>
<td>0.3 (0–12.5)</td>
<td>0.49</td>
</tr>
<tr>
<td>Leisure physical activity(^6)</td>
<td>60 (0–210)</td>
<td>40 (0–180)</td>
<td>30 (0–150)</td>
<td>0.70</td>
</tr>
<tr>
<td>Plasma folate (ng/mL)(^6)</td>
<td>14 (8–27)(^a)</td>
<td>16 (8–25)(^a)</td>
<td>20 (14–31)(^b)</td>
<td>0.002</td>
</tr>
<tr>
<td>RBC folate (ng/mL)(^6)</td>
<td>424 (321–771)(^a)</td>
<td>572 (361–852)(^b)</td>
<td>726 (431–863)(^b)</td>
<td>0.002</td>
</tr>
<tr>
<td>Plasma Hcy (µmol/L)(^6)</td>
<td>8.6 (6.9–10.8)</td>
<td>8.5 (6.8–9.7)</td>
<td>7.9 (6.8–9.7)</td>
<td>0.13</td>
</tr>
<tr>
<td>Plasma MMA (nmol/L)(^6)</td>
<td>158 (113–204)</td>
<td>146 (120–179)</td>
<td>168 (136–211)</td>
<td>0.17</td>
</tr>
<tr>
<td>Plasma vitamin B12 (pg/mL)(^6)</td>
<td>523 (354–703)</td>
<td>481 (352–650)</td>
<td>489 (371–685)</td>
<td>0.43</td>
</tr>
<tr>
<td>Plasma PLP (nmol/L)(^6)</td>
<td>62 (43–109)</td>
<td>70 (46–114)</td>
<td>68 (41–132)</td>
<td>0.83</td>
</tr>
<tr>
<td>Plasma choline (µmol/L)(^4)</td>
<td>9.2 ± 2.0</td>
<td>9.3 ± 2.2</td>
<td>9.4 ± 1.7</td>
<td>0.79</td>
</tr>
<tr>
<td>Plasma betaine (µmol/L)(^4)</td>
<td>28 ± 10</td>
<td>28 ± 10</td>
<td>25 ± 10</td>
<td>0.16</td>
</tr>
<tr>
<td>Plasma DMG (µmol/L)(^6)</td>
<td>2.5 (2.1–2.9)</td>
<td>2.3 (1.9–2.9)</td>
<td>2.4 (1.9–2.7)</td>
<td>0.07</td>
</tr>
<tr>
<td>Plasma TMAO (µmol/L)(^6)</td>
<td>3.8 (2.5–6.1)</td>
<td>3.7 (2.6–5.4)</td>
<td>4.1 (2.9–6.1)</td>
<td>0.33</td>
</tr>
<tr>
<td>Plasma creatinine (mg/dL)(^4)</td>
<td>0.72 ± 0.13</td>
<td>0.70 ± 0.11</td>
<td>0.73 ± 0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>DFE (µg/d)(^6)</td>
<td>409 (304–537)</td>
<td>411 (319–537)</td>
<td>441 (306–595)</td>
<td>0.51</td>
</tr>
<tr>
<td>Dietary vitamin B6 intake (mg/d)(^6)</td>
<td>1.4 (1.1–1.9)</td>
<td>1.4 (1.0–1.8)</td>
<td>1.5 (1.0–1.9)</td>
<td>0.43</td>
</tr>
<tr>
<td>Dietary vitamin B12 intake (µg/d)(^6)</td>
<td>5.1 (3.3–7.3)</td>
<td>4.1 (2.7–6.5)(^a)</td>
<td>5.3 (3.5–7.4)(^a)</td>
<td>0.01</td>
</tr>
<tr>
<td>Supplemental vitamin B2 intake (mg/d)(^6)</td>
<td>0.0 (0–1.7)</td>
<td>0.1 (0–1.7)</td>
<td>0.0 (0–1.7)</td>
<td>0.84</td>
</tr>
<tr>
<td>Supplemental vitamin B6 intake (mg/d)(^6)</td>
<td>0.0 (0–2.0)</td>
<td>1.0 (0–2.0)</td>
<td>0.0 (0–2.0)</td>
<td>0.32</td>
</tr>
<tr>
<td>Supplemental vitamin B12 intake (µg/d)(^6)</td>
<td>0.0 (0–6.0)</td>
<td>2.3 (0–6.0)</td>
<td>0.0 (0–6.0)</td>
<td>0.89</td>
</tr>
</tbody>
</table>

\(^1\)The study participants were a subset of the control group from a nested case-control study investigating colorectal cancer risk in the Women’s Health Initiative-Observational Study.

\(^2\)Differences between FA fortification periods were analyzed by chi-square tests (categorical variables), one-way ANOVA (normally distributed continuous variables), or non-parametric Kruskal-Wallis tests (non-normally distributed continuous variables); different superscript letters within a row indicate a difference between FA fortification periods at \(P < 0.05\).

\(^3\)\(n = 122\) in the pre-fortification period; \(n = 204\) in the peri-fortification period; \(n = 82\) in the post-fortification period.

\(^4\)Values are mean ± SD for normally distributed continuous variables.

\(^5\)Values are percentage for categorical variables.

\(^6\)Values are median (interquartile range) for non-normally distributed continuous variables.
African American, Hispanic, Asian or Pacific Islander, American Indian or Alaskan Native. Abbreviations used: BMI, body mass index; RBC, red blood cell; Hcy, homocysteine; MMA, methylmalonic acid; PLP, pyridoxal-5’-phosphate; DMG, dimethylglycine; TMAO, trimethylamine N-oxide; DFE, dietary folate equivalent.
Effect of FA fortification period on baseline leukocyte global DNA methylation

Leukocyte global DNA methylation did not differ \((P = 0.86,\) unadjusted; \(P = 0.38,\) multivariate-adjusted) among FA fortification periods (Table 1.2). However, we observed an interaction \((P = 0.02)\) between fortification period and RBC folate status in relation to DNA methylation. Specifically, the highest (vs. lowest) RBC folate group had higher marginal mean DNA methylation \((5.12\% vs. 4.99\%; P = 0.05)\) in the pre-fortification period, but lower DNA methylation \((4.95\% vs. 5.16\%; P = 0.03)\) in the post-fortification period (Table 1.3). In addition, leukocyte global DNA methylation tended to differ \((P = 0.08;\) multivariate-adjusted) among fortification periods \((post < peri < pre)\) within the highest RBC folate group, but not within the lowest RBC folate group \((P = 0.20;\) multivariate-adjusted) (Table 1.3).

**Table 1.2.** Baseline leukocyte global DNA methylation levels (%) according to folic acid (FA) fortification period

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n Value</td>
<td>n Value</td>
<td>n Value</td>
</tr>
<tr>
<td>Unadjusted (^2)</td>
<td>122 5.04 ± 0.35</td>
<td>204 5.05 ± 0.37</td>
<td>82 5.03 ± 0.38</td>
</tr>
<tr>
<td>Multivariate-adjusted (^3)</td>
<td>118 5.04 ± 0.03</td>
<td>199 5.06 ± 0.03</td>
<td>77 5.00 ± 0.04</td>
</tr>
</tbody>
</table>

\(^1\) Linear regression models were used to compare mean DNA methylation across fortification periods. 
\(^2\) Values are mean ± SD for unadjusted analyses. 
\(^3\) Multivariate analyses were adjusted for age, BMI, ethnicity, creatinine, and \(MTHFR\) C677T genotype; values are marginal mean ± SE.
Table 1.3. Baseline leukocyte global DNA methylation levels (%) within the lowest and highest RBC folate groups according to folic acid (FA) fortification period.\(^1,2\)

<table>
<thead>
<tr>
<th></th>
<th>Lowest RBC Folate group</th>
<th>Highest RBC Folate group</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n)</td>
<td>Value</td>
<td>(n)</td>
</tr>
<tr>
<td>Pre-fortification (1994–1995)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted(^3)</td>
<td>71</td>
<td>5.01 ± 0.37</td>
<td>51</td>
</tr>
<tr>
<td>Multivariate-adjusted(^4)</td>
<td>69</td>
<td>4.99 ± 0.04</td>
<td>49</td>
</tr>
<tr>
<td>Unadjusted(^3)</td>
<td>102</td>
<td>5.09 ± 0.38</td>
<td>102</td>
</tr>
<tr>
<td>Multivariate-adjusted(^4)</td>
<td>101</td>
<td>5.08 ± 0.04</td>
<td>98</td>
</tr>
<tr>
<td>Post-fortification (1998)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted(^3)</td>
<td>29</td>
<td>5.18 ± 0.39</td>
<td>53</td>
</tr>
<tr>
<td>Multivariate-adjusted(^4)</td>
<td>25</td>
<td>5.16 ± 0.07</td>
<td>52</td>
</tr>
</tbody>
</table>

\(^1\)Participants were divided into tertiles of RBC folate, and lowest (<471 ng/mL) and highest (>672 ng/mL) RBC folate groups were further stratified by FA fortification period.

\(^2\)Linear regression models were used to compare differences between RBC folate groups.

\(^3\)Values are mean ± SD for unadjusted analyses.

\(^4\)Multivariate analyses were adjusted for age, BMI, ethnicity, creatinine, and \(MTHFR\) C677T genotype; values are marginal mean ± SE.

Correlations between baseline leukocyte global DNA methylation and one-carbon biomarkers according to FA fortification period

The univariate Spearman correlations between one-carbon biomarkers and leukocyte global DNA methylation are shown in Table 1.4. Prior to fortification, there were significant, but modest, positive associations of global DNA methylation with plasma folate \((r = 0.20, P = 0.04)\) and RBC folate \((r = 0.24, P = 0.01)\) as well as a borderline significant positive association with plasma vitamin B12 \((r = 0.18, P = 0.06)\). Global DNA methylation was also inversely correlated with plasma methylmalonic acid (MMA; \(r = -0.26, P = 0.03\)), choline \((r = -0.31, P = 0.002)\) and homocysteine \((r = -0.26, P = 0.007)\). In the peri-fortification period, no significant relationships were observed between one-carbon biomarkers and global DNA methylation. Finally, in the post-fortification period, global DNA methylation was positively correlated with plasma homocysteine \((r = 0.28, P = 0.02)\).
Table 1.4. Spearman rank correlation coefficients (r) between baseline leukocyte global DNA methylation and one-carbon biomarkers according to folic acid (FA) fortification period

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>r</td>
<td>P value</td>
</tr>
<tr>
<td>Plasma folate (ng/mL)</td>
<td>115</td>
<td>0.20</td>
<td>0.04</td>
</tr>
<tr>
<td>RBC folate (ng/mL)</td>
<td>118</td>
<td>0.24</td>
<td>0.01</td>
</tr>
<tr>
<td>Plasma vitamin B12 (pg/mL)</td>
<td>115</td>
<td>0.18</td>
<td>0.06</td>
</tr>
<tr>
<td>Plasma MMA (nmol/L)</td>
<td>74</td>
<td>−0.26</td>
<td>0.03</td>
</tr>
<tr>
<td>Plasma choline (µmol/L)</td>
<td>102</td>
<td>−0.31</td>
<td>0.002</td>
</tr>
<tr>
<td>Plasma betaine (µmol/L)</td>
<td>102</td>
<td>−0.09</td>
<td>0.39</td>
</tr>
<tr>
<td>Plasma DMG (µmol/L)</td>
<td>102</td>
<td>−0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>Plasma TMAO (µmol/L)</td>
<td>102</td>
<td>0.001</td>
<td>0.99</td>
</tr>
<tr>
<td>Plasma Hcy (µmol/L)</td>
<td>115</td>
<td>−0.26</td>
<td>0.007</td>
</tr>
<tr>
<td>Plasma cysteine (µmol/L)</td>
<td>115</td>
<td>−0.10</td>
<td>0.28</td>
</tr>
</tbody>
</table>

1Analyses were adjusted for age, BMI, ethnicity, creatinine, and MTHFR C677T genotype. Abbreviations used: RBC, red blood cell; MMA, methylmalonic acid; DMG, dimethylglycine; TMAO, trimethylamine N-oxide; Hcy, homocysteine.

Main predictors of baseline leukocyte global DNA methylation according to FA fortification period

One-carbon biomarkers that predicted global DNA methylation were identified according to FA fortification period, testing them individually in multivariate-adjusted models (Table 1.5). Prior to fortification, RBC folate positively predicted global DNA methylation (β = 0.25, P = 0.02) explaining 5% of the residual variation (partial $R^2 = 0.05$). Plasma vitamin B12 also tended to positively predict DNA methylation (β = 0.20, P = 0.08) explaining 3% of the residual variation (partial $R^2 = 0.03$); however, plasma homocysteine (β = -21.23, P = 0.03), MMA (β = -1.18, P = 0.05), and choline (β = -57.82, P = 0.002) negatively predicted global DNA methylation explaining 4% (partial $R^2 = 0.04$), 6% (partial $R^2 = 0.06$) and 10% (partial $R^2 = 0.10$) of the residual variation, respectively. In the peri-fortification period, no significant predictors of
DNA methylation were detected. Finally, in the post-fortification period, plasma homocysteine tended to positively predict global DNA methylation ($\beta = 29.37$, $P = 0.07$) explaining 4% of the residual variation (partial $R^2 = 0.04$). The overall $R^2$ explained by one-carbon biomarkers, tested in an unadjusted linear regression model in which all variables were included simultaneously, was 0.12 in the pre- and peri-fortification periods and 0.19 in the post-fortification period (Supplementary Table S1.1) with plasma choline, plasma dimethylglycine (DMG), and plasma homocysteine being the strongest predictors in each period, respectively (Supplementary Table S1.2).

Table 1.5. Predictors of baseline leukocyte global DNA methylation according to folic acid (FA) fortification period$^{1,2}$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$\beta$ Coefficient</td>
<td>$P$ value</td>
</tr>
<tr>
<td>Plasma folate (ng/mL)</td>
<td>115</td>
<td>3.98</td>
<td>0.31</td>
</tr>
<tr>
<td>RBC folate (ng/mL)</td>
<td>118</td>
<td>0.25</td>
<td>0.02</td>
</tr>
<tr>
<td>Plasma vitamin B12 (pg/mL)</td>
<td>115</td>
<td>0.20</td>
<td>0.08</td>
</tr>
<tr>
<td>Plasma MMA (nmol/L)</td>
<td>74</td>
<td>−1.18</td>
<td>0.05</td>
</tr>
<tr>
<td>Plasma choline (µmol/L)</td>
<td>102</td>
<td>−57.82</td>
<td>0.002</td>
</tr>
<tr>
<td>Plasma betaine (µmol/L)</td>
<td>102</td>
<td>−2.91</td>
<td>0.41</td>
</tr>
<tr>
<td>Plasma DMG (µmol/L)</td>
<td>102</td>
<td>−31.11</td>
<td>0.43</td>
</tr>
<tr>
<td>Plasma TMAO (µmol/L)</td>
<td>102</td>
<td>6.10</td>
<td>0.12</td>
</tr>
<tr>
<td>Plasma Hcy (µmol/L)</td>
<td>115</td>
<td>−21.23</td>
<td>0.03</td>
</tr>
<tr>
<td>Plasma cysteine (µmol/L)</td>
<td>115</td>
<td>−1.23</td>
<td>0.19</td>
</tr>
</tbody>
</table>

$^1$Linear regression models were used, adjusting for age, BMI, ethnicity, creatinine, and $MTHFR$ C677T genotype.

$^2$Beta ($\beta$) coefficient indicates mean increase in DNA methylation per 1000-unit increase in one-carbon biomarker. Abbreviations used: RBC, red blood cell; MMA, methylmalonic acid; DMG, dimethylglycine; TMAO, trimethylamine N-oxide; Hcy, homocysteine.
DISCUSSION

The present study investigated the association between mandatory FA fortification and leukocyte global DNA methylation, as well as the relationship between global DNA methylation, RBC folate, and other biomarkers of one-carbon metabolism in postmenopausal women. The following two main findings emerged: (1) FA fortification period and RBC folate status interacted to influence global DNA methylation and; (2) associations between one-carbon biomarkers and global DNA methylation differed between FA fortification periods.

FA fortification period interacted with RBC folate status to influence global DNA methylation

Previous studies have found that global DNA methylation can be altered by folate depletion or repletion in healthy adults.\textsuperscript{14-17} In postmenopausal women, global DNA methylation significantly decreased under folate depletion\textsuperscript{14,15} and increased upon folate repletion.\textsuperscript{14} Based on these findings and the role of folate as a methyl donor, we anticipated that global DNA methylation would be higher among postmenopausal women with higher (vs. lower) RBC folate status. This expected result was observed in the pre-fortification period, but not in the post-fortification period during which women with higher (vs. lower) RBC folate status had lower DNA methylation.

Excess FA intake through fortified foods and supplements can lead to the accumulation of unmetabolized FA,\textsuperscript{18} which may interfere with normal folate metabolism\textsuperscript{19-22} and lower global DNA methylation.\textsuperscript{23} Although supraphysiologic folate status (i.e., total plasma folate concentrations $> 19.8$ ng/mL)\textsuperscript{24} was observed among postmenopausal women with higher RBC folate in the post-fortification period (25.3 ng/mL; Supplementary Table S1.3), it was similarly observed among women with higher RBC folate in the pre-fortification period (25.1 ng/mL) and
appeared to be mostly attributable to FA supplement use in both periods. Specifically, a higher percentage of FA supplement users was observed among participants with higher (vs. lower) RBC concentrations (76% vs. 20%; Supplementary Table S1.4), and higher RBC folate concentrations were observed among FA supplement users (vs. non-users) across fortification periods (Supplementary Table S1.5). Nonetheless, a previous study conducted in the US reported the highest concentration of plasma unmetabolized FA in subjects exposed to both FA fortified foods and supplements as compared to those exposed only to FA fortified foods, or only to supplements. Thus, it is possible that unmetabolized FA was elevated to a greater extent in the post- (vs. pre-) fortification period among women with higher RBC folate status. Measurements of unmetabolized FA in our cohort are needed to further explore this possibility, and additional studies are required to clarify the health outcomes, if any, of the inverse relationship between leukocyte global DNA methylation and high RBC folate in the era of FA fortification.

**Associations between one-carbon biomarkers and global DNA methylation differed among FA fortification periods**

Previous human studies have reported conflicting results with positive or no relationships between circulating folate (i.e., plasma and RBC folate) and global DNA methylation. In the present study, plasma and RBC folate were positively correlated with DNA methylation in the pre-fortification period, but not in the peri- or post-fortification period. These findings suggest that the relationship between folate status and global DNA methylation is nonlinear and that folate status is likely to be a stronger predictor of global DNA methylation when folate availability is lower (i.e., prior to FA fortification). However, as alluded to above, it is possible that the differences in the relationship between circulating folate and DNA
methylation across fortification periods arose from differences in the amounts of metabolized and unmetabolized folate. For example, metabolized folate present in the pre-fortification period may positively associate with global DNA methylation, while unmetabolized FA more likely to be present in the peri- and post-fortification periods may attenuate the positive relationship between folate status and global DNA methylation. Taken together, when total folate status (metabolized plus unmetabolized) is considered across the full spectrum from deficiency to very high, the overall association between folate and global DNA methylation may approximate a reverse U-shaped curve rather than a linear relationship.

Folate intake/status may also modify the relationship between DNA methylation and other nutrients involved in one-carbon metabolism. Indeed, biomarkers of vitamin B12 status (i.e., plasma vitamin B12 and MMA) were associated with leukocyte global DNA methylation in the pre-fortification period, but not in the peri- or post-fortification period. Both folate and vitamin B12 are required for the provision of methyl groups through the methionine synthase reaction (Supplementary Figure S1.1). However, folate is suggested to be a stronger determinant of biomarkers of the methylation cycle (e.g., plasma homocysteine) than vitamin B12, which may explain the lack of association between vitamin B12 status and global DNA methylation in the peri- and post-fortification periods.

The relationship between plasma choline and global DNA methylation was also modified by FA exposure with an inverse relationship observed in the pre-fortification period, but not in the peri- or post-fortification period. The inverse relationship between choline (a methyl donor) and DNA methylation in the pre-fortification period is unexpected and requires confirmation in other studies. However, when folate is less abundant (i.e., prior to FA fortification), supply of S-adenosylmethionine (SAM) for methylation reactions may be reduced thereby creating a
competition among the various methyltransferases. As the affinity of DNA methyltransferase for SAM is \(~18\) times higher than phosphatidylethanolamine \(N\)-methyltransferase,\(^{32}\) the enzyme that produces choline endogenously, SAM may be preferentially partitioned toward DNA methylation thus reducing endogenous choline production. In turn, this could lead to the inverse relationship observed in the pre-fortification period between DNA methylation and plasma choline.

Prior to FA fortification, we observed an inverse relationship between plasma homocysteine and global DNA methylation, which is consistent with previous reports.\(^{26,33}\) Interestingly, however, plasma homocysteine was positively correlated with DNA methylation in the post-fortification period. The divergent relationships between homocysteine and DNA methylation across fortification periods may arise from the fact that homocysteine is both a precursor and product of cellular methylation reactions. These data collectively suggest that the relationship between homocysteine and DNA methylation is dynamic and likely to be dependent on folate availability.

**Strengths and limitations**

The present study had several strengths including: (1) a unique opportunity to investigate the impact of mandatory FA fortification on global DNA methylation by stratifying into three fortification periods (pre-, peri-, and post-) and; (2) examination of a wide range of biomarkers involved in one-carbon metabolism as potential predictors of global DNA methylation according to FA fortification period. Several limitations should also be noted: (1) relatively small sample size; (2) potential for residual confounding by factors that were either not collected in the WHI-OS or not measured with sufficient precision and; (3) single measures of one-carbon biomarkers
and global DNA methylation within each FA fortification period, which may not fully reflect the true complexity of DNA methylation reactions.

**Conclusion**

These data suggest that during FA fortification, higher RBC folate status is associated with a reduction in leukocyte global DNA methylation among postmenopausal women. If reductions in leukocyte global DNA methylation are shown to have adverse health outcomes in future studies, FA supplement use may not be advisable among postmenopausal women residing in the US or other countries with mandated FA fortification programs. The present study also suggests that FA intake via fortification modifies the relationship between one-carbon biomarkers and global DNA methylation, but potential biologic mechanisms need discerning.
MATERIALS AND METHODS

Subjects and study design

The WHI-OS is a prospective cohort study that was established to investigate the predictors and causes of morbidity and mortality in postmenopausal women.\textsuperscript{34,35} The study enrolled 93676 postmenopausal women, aged 50–79 y, at 40 clinical Centers throughout the US between 1993 and 1998. These years of enrollment spanned the pre- (prior to January 1, 1996), peri- (1996–1997), and post- (after January 1, 1998) FA fortification periods in the US.\textsuperscript{9} Women were excluded from the study if they had medical conditions with a predicted survival of less than 3 y; if they had adherence/retention issues (alcoholism, drug dependency, mental illness, or dementia); or if they were participating in another clinical trial. The study was approved by the human subject review boards at the Fred Hutchinson Cancer Research Center where the WHI Clinical Coordinating Center is located and at all 40 clinical centers. Written informed consent was obtained from all participants.\textsuperscript{34,35}

In the present study, participants were a subset of those from a nested case-control study investigating colorectal cancer risk in the WHI-OS.\textsuperscript{12,13} From the controls of the study, we selected 408 women from the lowest (n = 202) and highest (n = 206) tertiles of baseline RBC folate concentrations, matching on age and timing of the baseline blood draw, which spanned the following FA fortification periods: pre-fortification (1994–1995; n = 71 low tertile, 51 high tertile), peri-fortification (1996–1997; n = 102 low tertile, 102 high tertile), and post-fortification (1998; n = 29 low tertile, 53 high tertile). The proportions in the fortification periods correspond approximately to the recruitment of the WHI-OS. However, because the nested case-control study from which we were sampling did not contain at least 50 participants in the post-fortification period, low RBC folate group, we selected additional participants from the pre-
fortification, low RBC folate group in order to maintain approximately the same number of participants in each of the low and high RBC folate tertiles.

Data collection

Baseline demographic and health-related characteristics (i.e., age, race/ethnicity, education, smoking status, and physical activity) were collected using standardized questionnaires. Height and weight were measured using a standardized protocol, and BMI was calculated as weight (kg)/height (m$^2$). Dietary intake of folate, vitamin B6 and vitamin B12 was based on data derived from the WHI food-frequency questionnaire as previously described. Supplemental vitamin intakes of B2, B6, and B12 were assessed by an inventory in which nutrients were recorded based on participants’ current dietary supplement bottles, which they brought to the clinic visits. To account for differences in bioavailability between synthetic FA and natural food folate, dietary folate equivalent (DFE) was used as the unit for total folate intake.

Analytic measurements

Blood samples were drawn at baseline after at least 12 h of fasting. Samples were kept at 4 °C for up to 1 h prior to centrifugation. Plasma and serum were collected and stored at -70 °C until analysis. Leukocyte global DNA methylation was measured in de-identified samples using liquid chromatography-tandem mass spectrometry (LC-MS/MS) as described by Song et al. with modifications based on our instrumentation. A total of 11 batches (40 samples per batch) were run in duplicate. For each batch, all samples from both comparisons (i.e., high and low RBC folate) were randomly ordered and equally represented and matched on period of blood
draw by pre-, peri- and post-fortification and on age. Both internal laboratory controls and 10% blind duplicate samples were used to determine assay precision and monitor assay performance. Internal laboratory controls included: (1) unmethylated lambda DNA (Promega); (2) methylated lambda DNA (~30% of DNA methylated); (3) four in-house human biological control samples and; (4) a negative control (water). All quality control samples were prepared in duplicate and interspersed among the samples. DNA methylation is expressed as a percentage of total cytosine: 

\[
\text{[methylated cytosine} / (\text{methylated} + \text{unmethylated cytosine}) \times 100\%.
\]

Plasma concentrations of choline and its metabolites (i.e., betaine, DMG, trimethylamine N-oxide [TMAO]) were measured using stable isotope dilution LC-MS/MS methodology.\(^4\) Plasma total homocysteine and cysteine were determined by high-pressure liquid chromatography (HPLC) with post-column fluorescence detection\(^4\); plasma and RBC folate as well as plasma vitamin B12 were measured by radioassay (SimulTRAC; MP Biomedicals); plasma pyridoxal-5’-phosphate (PLP) was analyzed by HPLC with fluorescence detection\(^4\); plasma MMA was measured by LC-MS/MS\(^4\); plasma creatinine was quantified by the Jaffe rate reaction method (DxC Instrument; Beckman Coulter); and methylenetetrahydrofolate reductase (MTHFR) C677T genotype (rs 1801133) was determined by the Illumina 384-plex BeadXpress genotyping platform (Illumina Inc.).

Inter-assay coefficients of variance of the blind duplicate control samples for each of the assays were as follows: global DNA methylation, 5.5%; choline, 5.6%; betaine, 4.6%; DMG, 11.9%; TMAO, 5.8%; homocysteine, 6.5%; cysteine, 7.1%; RBC folate, 10.2%; plasma folate, 4.8%; vitamin B12, 6.2%; PLP, 5.9%; MMA, 15.0%; and creatinine, 4.1%.
Statistical analysis

Differences in baseline characteristics of the study population between FA fortification periods were analyzed by: (1) one-way analysis of variance (ANOVA) for normally distributed continuous variables; (2) non-parametric Kruskal-Wallis tests for non-normally distributed continuous variables or; (3) chi-square tests for categorical variables. Linear regression models were used to: (1) examine the influence of FA fortification and RBC folate (and their interaction term) on baseline leukocyte global DNA methylation and; (2) identify the one-carbon biomarkers that predicted leukocyte global DNA methylation. The partial $R^2$ for each predictor variable was determined to estimate the contribution of each predictor to the total variability in DNA methylation. Spearman rank correlation coefficients ($r$) were also computed to examine associations between leukocyte global DNA methylation and one-carbon biomarkers. In the multivariate-adjusted analyses, we controlled for age and BMI along with plasma creatinine, ethnicity (white/not white) and $MTHFR$ C677T genotype as these three variables were shown to be influential in a univariate model assessing possible confounders on DNA methylation. $MTHFR$ C677T genotype was treated as an additive variable (i.e., minor allele count) in our statistical models because of reduced variation; parameter estimates were not changed substantially when $MTHFR$ C677T genotype was treated as a categorical variable. Because approximately 25% of plasma creatinine values were missing among the sample due to insufficient sample availability, simple mean imputation was used for the missing creatinine values. Model results using multiple imputation and simple mean imputation were similar. Significance was defined as $P < 0.05$, and all statistical tests were 2-sided. The data were analyzed by SAS version 9.3 (SAS Institute Inc.).
REFERENCES


**Supplementary Figure S1.1** Simplified diagram of folate- and choline-mediated DNA methylation reactions. Relevant enzymes are highlighted in gray. Abbreviations: BHMT, betaine homocysteine methyltransferase; DHF, dihydrofolate; DHFR, dihydrofolate reductase; DMG, dimethylglycine; DNMT, DNA methyltransferase; MS, methionine synthase; MTHFR, methylenetetrahydrofolate reductase; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine; THF, tetrahydrofolate.
### Supplementary Table S1.1 Overall $R^2$ explained by all one-carbon biomarkers according to folic acid (FA) fortification period$^{1,2}$

<table>
<thead>
<tr>
<th></th>
<th>Overall unadjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-fortification (1994-1995)</td>
<td>0.12</td>
</tr>
<tr>
<td>Peri-fortification (1996-1997)</td>
<td>0.12</td>
</tr>
<tr>
<td>Post-fortification (1998)</td>
<td>0.19</td>
</tr>
</tbody>
</table>

$^{1}$Unadjusted linear regression models were used by including all variables simultaneously.

$^{2}$n=73 in the pre-fortification period; n=123 in the peri-fortification period; n=45 in the post-fortification period.
**Supplementary Table S1.2.** Predictors of baseline leukocyte global DNA methylation according to folic acid (FA) fortification period with all variables included in the statistical model\(^1,2,3\)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\beta) Coefficient</td>
<td>(P) value</td>
<td>Partial (R^2)</td>
</tr>
<tr>
<td>Plasma folate (ng/mL)</td>
<td>-1.16</td>
<td>0.79</td>
<td>0.001</td>
</tr>
<tr>
<td>RBC folate (ng/mL)</td>
<td>0.11</td>
<td>0.53</td>
<td>0.006</td>
</tr>
<tr>
<td>Plasma vitamin B12 (pg/mL)</td>
<td>0.04</td>
<td>0.84</td>
<td>0.001</td>
</tr>
<tr>
<td>Plasma MMA (nmol/L)</td>
<td>-0.49</td>
<td>0.38</td>
<td>0.012</td>
</tr>
<tr>
<td>Plasma choline (µmol/L)</td>
<td>-44.37</td>
<td>0.06</td>
<td>0.057</td>
</tr>
<tr>
<td>Plasma betaine (µmol/L)</td>
<td>-1.78</td>
<td>0.75</td>
<td>0.002</td>
</tr>
<tr>
<td>Plasma DMG (µmol/L)</td>
<td>-0.85</td>
<td>0.99</td>
<td>0.000</td>
</tr>
<tr>
<td>Plasma TMAO (µmol/L)</td>
<td>7.10</td>
<td>0.31</td>
<td>0.016</td>
</tr>
<tr>
<td>Plasma Hcy (µmol/L)</td>
<td>16.22</td>
<td>0.51</td>
<td>0.007</td>
</tr>
<tr>
<td>Plasma cysteine (µmol/L)</td>
<td>0.08</td>
<td>0.96</td>
<td>0.000</td>
</tr>
</tbody>
</table>

\(^1\)Unadjusted linear regression models were used by including all variables simultaneously.

\(^2\)Beta (\(\beta\)) coefficient indicates mean increase in DNA methylation per 1000-unit increase in one-carbon biomarker.

\(^3\)n=73 in the pre-fortification period; n=123 in the peri-fortification period; n=45 in the post-fortification period.

Abbreviations used: RBC, red blood cell; MMA, methylmalonic acid; DMG, dimethylglycine; TMAO, trimethylamine N-oxide; Hcy, homocysteine.
**Supplementary Table S1.3.** Plasma folate concentrations (ng/mL) within the lowest and highest RBC folate groups according to folic acid (FA) fortification period\(^1\)

<table>
<thead>
<tr>
<th></th>
<th>Lowest RBC Folate group</th>
<th>Highest RBC Folate group</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Median (IQR)</td>
<td>n</td>
</tr>
<tr>
<td>Pre-fortification (1994-1995)</td>
<td>70</td>
<td>8.5 (5.8-14.6)</td>
<td>49</td>
</tr>
<tr>
<td>Peri-fortification (1996-1997)</td>
<td>100</td>
<td>9.2 (5.8-13.7)</td>
<td>100</td>
</tr>
<tr>
<td>Post-fortification (1998)</td>
<td>28</td>
<td>12.8 (9.7-20.0)</td>
<td>51</td>
</tr>
</tbody>
</table>

\(^1\)Linear regression models were used to compare median plasma folate concentrations between RBC folate groups and were adjusted for age, BMI, ethnicity, creatinine, and \(MTHFR\) C677T genotype.
Supplementary Table S1.4. Folic acid (FA) supplement use (%) within the lowest (n=202) and highest (n=206) RBC folate groups.

<table>
<thead>
<tr>
<th>FA supplement use</th>
<th>Lowest RBC folate group</th>
<th>Highest RBC folate group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>19.8% (40/202)</td>
<td>75.7% (156/206)</td>
</tr>
<tr>
<td>No</td>
<td>80.2% (162/202)</td>
<td>24.3% (50/206)</td>
</tr>
</tbody>
</table>
**Supplementary Table S1.5.** RBC folate concentrations (ng/mL) within folic acid (FA) supplement users and non-supplement users according to FA fortification period¹

<table>
<thead>
<tr>
<th></th>
<th>FA supplement users</th>
<th>Non-FA supplement users</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean ± SD</td>
<td>n</td>
</tr>
<tr>
<td>Pre-fortification (1994-1995)</td>
<td>56</td>
<td>761 ± 282</td>
<td>66</td>
</tr>
<tr>
<td>Peri-fortification (1996-1997)</td>
<td>103</td>
<td>820 ± 317</td>
<td>101</td>
</tr>
<tr>
<td>Post-fortification (1998)</td>
<td>37</td>
<td>814 ± 276</td>
<td>45</td>
</tr>
</tbody>
</table>

¹Linear regression models were used to compare mean RBC folate concentrations between FA supplement users and non-supplement users and were adjusted for age, BMI, ethnicity, creatinine, and *MTHFR* C677T genotype.
CHAPTER 2
Plasma Choline Metabolites and Colorectal Cancer Risk in the Women's Health Initiative Observational Study*

ABSTRACT

Few studies have examined associations between plasma choline metabolites and risk of colorectal cancer. Therefore, we investigated associations between plasma biomarkers of choline metabolism [choline, betaine, dimethylglycine, and trimethylamine N-oxide (TMAO)] and colorectal cancer risk among postmenopausal women in a case–control study nested within the Women’s Health Initiative Observational Study. We selected 835 matched case–control pairs, and cases were further stratified by tumor site (proximal, distal, or rectal) and stage (local/regional or metastatic). Colorectal cancer was assessed by self-report and confirmed by medical records over the mean of 5.2 years of follow-up. Baseline plasma choline metabolites were measured by LC/MS-MS. In multivariable-adjusted conditional logistic regression models, plasma choline tended to be positively associated with rectal cancer risk [OR (95% confidence interval, CI) highest vs. lowest quartile = 2.44 (0.93–6.40); \( P \) trend = 0.08], whereas plasma betaine was inversely associated with colorectal cancer overall [0.68 (0.47–0.99); \( P \) trend = 0.01] and with local/regional tumors [0.64 (0.42–0.99); \( P \) trend = 0.009]. Notably, the plasma betaine:choline ratio was inversely associated with colorectal cancer overall [0.56 (0.39–0.82); \( P \) trend = 0.004] as well as with proximal [0.66 (0.41–1.06); \( P \) trend = 0.049], rectal [0.27 (0.10–0.78); \( P \) trend = 0.02], and local/regional [0.50 (0.33–0.76); \( P \) trend = 0.001] tumors. Finally, plasma TMAO, an oxidative derivative of choline produced by intestinal bacteria, was positively associated with rectal cancer [3.38 (1.25–9.16); \( P \) trend = 0.02] and with overall colorectal cancer risk among women with lower (vs. higher) plasma vitamin B12 levels (\( P \) interaction = 0.003). Collectively, these data suggest that alterations in choline metabolism, which may arise early in disease development, may be associated with higher risk of colorectal cancer. The positive association between plasma TMAO and colorectal cancer risk is consistent with an involvement of the gut.
microbiome in colorectal cancer pathogenesis.
INTRODUCTION

Colorectal cancer is the third most commonly diagnosed cancer in both men and women and a major cause of cancer deaths in the United States (1). Disturbances in one-carbon metabolism, which lead to genomic instability (e.g., aberrant DNA methylation and DNA damage), may contribute to colorectal cancer development (2, 3). Choline and folate are methyl nutrients involved in one-carbon metabolism and play a critical role in methylation reactions, including DNA methylation, as well as DNA stability and repair (4–6). Although low folate intake and low circulating levels of folate are associated with high risk of colorectal cancer (2, 7–9), less is known about the association between choline and colorectal cancer risk.

Choline participates in methylation reactions following its oxidation to betaine, which donates a methyl group for homocysteine remethylation, forming methionine and dimethylglycine (DMG). Betaine also serves as an osmolyte and plays a major role in protecting cells from hyperosmotic stress that can lead to chronic inflammation, a risk factor for colorectal cancer (1, 10, 11). To date, only a few studies have examined the association between plasma betaine and colorectal carcinogenesis. In a Norwegian population, plasma betaine was inversely associated with the occurrence of distal colorectal adenomas (12). A recent case–control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) also reported an inverse association between plasma betaine and colorectal cancer risk among participants with low plasma folate concentrations (13).

Choline can also undergo catabolism by the intestinal bacteria to form trimethylamine (TMA), which is further converted to trimethylamine N-oxide (TMAO) by the liver enzyme flavin monooxygenase (FMO; refs. 14, 15). Although intestinal microbiota have been implicated in the development of colorectal cancer (16–18), the association between gut microbiota–
dependent choline metabolites and colorectal cancer risk is unknown.

In this report, we investigated the associations between plasma biomarkers of choline metabolism (choline, betaine, DMG, and TMAO) and colorectal cancer risk in a case–control study nested within the Women's Health Initiative Observational Study (WHI-OS) cohort. Because of the interdependence of choline and folate as well as other B vitamins (vitamin B6 and B12) in one-carbon metabolism (4, 5), we further explored their influence and that of folic acid (FA) fortification (19) on the associations between plasma choline metabolites and colorectal cancer risk.
Patients and Methods

Study population

The WHI-OS is a prospective cohort study designed to investigate the predictors and causes of morbidity and mortality in postmenopausal women (20, 21). The study enrolled 93,676 postmenopausal women, ages 50 to 79 years, at 40 centers throughout the United States between 1993 and 1998. Women were excluded if they had medical conditions with a predicted survival of <3 years; if they had adherence/retention issues; or if they were participating in another clinical trial.

For the present study, incident colorectal cancer cases were selected as of April 24, 2008, and the average time from baseline to colorectal cancer diagnosis was 5.2 ± 3.1 years (mean ± SD; refs. 11, 22). Women were excluded if they had a history of colorectal cancer or in situ colorectal cancer; if they had no available biospecimens; or if a death certificate provided the only report of colorectal cancer. Controls who were free of cancer at the time of case diagnosis were selected from the WHI-OS by using risk-set sampling. Cases and controls were matched on age (± 3 years), race/ethnicity, timing of baseline blood draw (± 6 months), enrollment date (± 1 year), and baseline hysterectomy status (11, 22). Thus, the present study included 835 incident colorectal cancer cases and 835 matched controls. Approval for conducting the study was obtained from human subject review committees at the Fred Hutchinson Cancer Research Center (WHI Clinical Coordinating Center), as well as at all 40 clinical centers. Written informed consent was obtained from all participants.

Data collection

Demographic and health-related characteristics were collected at baseline using
standardized questionnaires (20). Height and weight were measured using a standardized protocol, and body mass index (BMI) was calculated as weight (kg)/height (m$^2$). Colorectal cancer was annually assessed using self-administered questionnaires collected from each participant by mail and during an in-person clinical follow-up visit at year 3 (23). All colorectal cancer cases were confirmed by physician adjudicators. The International Classification of Diseases for Oncology, second edition codes were used to identify colorectal cancer cases based on tumor site as previously described (11). The Surveillance Epidemiology and End Results (SEER) program guidelines of the NCI were used for classifications of cancer cases (23).

**Analytic measurements**

Blood samples were drawn at study baseline after at least 12 hours of fasting. Samples were kept at 4°C for up to 1 hour before centrifugation. Plasma and serum were collected and stored at -70°C until analysis (22). Plasma concentrations of choline and its metabolites (betaine, DMG, and TMAO) were measured in de-identified samples using LC/MS-MS methodology with modifications based on our instrumentation (24). Plasma and red blood cell (RBC) folate as well as plasma vitamin B12 were measured by radioassays (SimulTRAC; MP Biomedicals); plasma pyridoxal-5’-phosphate (PLP) was analyzed by high-pressure liquid chromatography (HPLC) with fluorescence detection (25); and total plasma homocysteine was determined by HPLC with postcolumn fluorescence detection (26). Interassay coefficients of variance of the blind duplicate control samples for each of the assays were as follows: choline, 7%; betaine, 5%; DMG, 9%; TMAO, 6%; plasma folate, 5%; RBC folate, 10%; vitamin B12, 6%; PLP, 6%; and homocysteine, 7%.
**Statistical analysis**

Baseline characteristics of colorectal cancer cases and controls were compared using (i) *t* tests for normally distributed continuous variables; (ii) Wilcoxon tests for non-normally distributed continuous variables; and (iii) $\chi^2$ tests for categorical variables. Associations among plasma concentrations of choline metabolites were assessed using Spearman correlation analysis. Plasma choline metabolites were divided into quartiles based on the distribution of the controls. Conditional logistic regression models were used to estimate ORs and 95% confidence intervals (CI) of colorectal cancer risk among quartiles of choline metabolites, using the lowest quartiles as reference groups. Because risk-set sampling was used for selecting matched controls, the conditional ORs yielded estimates of the incidence rate ratio in a full cohort study. We further explored the associations between the ratios of choline metabolites (i.e., betaine:choline, DMG:choline, and DMG: betaine) and colorectal cancer risk, because the ratios of these metabolites (vs. individual metabolite alone) are suggested to be stronger predictors of metabolic disturbances (27). The models were first adjusted only for age (continuous) and then further adjusted for baseline confounding factors selected a priori: BMI, pack-years of smoking, physical activity, use of postmenopausal hormone therapy, history of colonoscopy, RBC folate, plasma vitamin B12, PLP, and homocysteine. All of these factors were added in the model as continuous variables except for postmenopausal hormone therapy use (categorical: never, past, or current). Tests of linear trend across increasing quartiles of choline metabolites were conducted by the Wald test, using the median value for each quartile as a single continuous variable.

To explore whether the associations between choline metabolites and colorectal cancer risk were modified by B vitamins involved in one-carbon metabolism, we conducted analyses
stratified into high/low plasma concentrations of folate, PLP, and vitamin B12 based on median values among controls. We also examined the influence of FA fortification by stratifying into the following FA fortification periods based on the timing of baseline blood draw: pretreatment (1994–1995), peritreatment (1996–1997; when initial fortification began, but was not yet mandated), and posttreatment (1998; ref. 28). The Wald test was used to evaluate the effect modification including a two-way interaction term between the ordinal trend variables (choline metabolites) and effect modifiers (B vitamins or FA fortification period). Because the matching was broken, unconditional multiple logistic regression models were used in these stratified analyses, further adjusting for days to colorectal cancer diagnosis and ethnicity. Significance was defined as \( P < 0.05 \), and all statistical tests were two-sided. Analyses were conducted by SAS version 9.3 (SAS Institute Inc.).
Results

Characteristics of the study population

Baseline characteristics of the colorectal cancer cases and controls are shown in Table 2.1. Compared with the controls, the cases had a higher BMI, a greater number of cigarettes smoked among current smokers, fewer weekly minutes of moderate or strenuous physical activity, and had a different distribution pattern of postmenopausal hormone therapy use. The colorectal cancer group also had a lower percentage of previous colonoscopy, but a higher percentage of having history of a colon polyp removed.

Plasma choline, betaine, and DMG concentrations did not differ between cases and controls (Table 2.1). However, the cases (vs. controls) had higher ($P = 0.005$) median plasma concentrations of TMAO (4.0 vs. 3.8 µmol/L) and tended to have a lower ($P = 0.07$) mean plasma betaine:choline ratio (2.9 vs. 3.0). In addition, the cases had lower median plasma folate, PLP, and vitamin B12 as well as higher median plasma homocysteine.

Among the cases, tumors were classified by tumor site (proximal, distal, or rectal) and stage (local/regional or metastatic). More than half (59%; $n = 489$) of the tumors were proximal followed by distal (21%; $n = 177$) and rectal (19%; $n = 155$). Two percent ($n = 14$) of the tumors were not classified by tumor site because they were unknown or had overlapping lesions. In addition, when stratified by tumor stage, the majority of the cases (85%; $n = 712$) had localized or regional tumors, whereas 12% of the cases ($n = 104$) had distant metastases. Two percent ($n = 18$) of the tumors were not stratified by tumor stage because their stages were unknown or not determined.
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cases</th>
<th>Controls</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>835</td>
<td>835</td>
<td>0.52</td>
</tr>
<tr>
<td>BMI (kg/m&lt;sup&gt;2&lt;/sup&gt;)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>824</td>
<td>827</td>
<td>0.004</td>
</tr>
<tr>
<td>Race/ethnicity&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>White</td>
<td>711</td>
<td>711</td>
<td></td>
</tr>
<tr>
<td>Other&lt;sup&gt;e&lt;/sup&gt;</td>
<td>124</td>
<td>124</td>
<td></td>
</tr>
<tr>
<td>Family income ($)&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>0.30</td>
</tr>
<tr>
<td>&lt; 34,999</td>
<td>374</td>
<td>351</td>
<td></td>
</tr>
<tr>
<td>35,000–74,999</td>
<td>294</td>
<td>282</td>
<td></td>
</tr>
<tr>
<td>≥ 75,000</td>
<td>111</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>Do not know</td>
<td>22</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Education (high school or less)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>160</td>
<td>186</td>
<td>0.11</td>
</tr>
<tr>
<td>Residence location (US region)&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>0.57</td>
</tr>
<tr>
<td>Northeast</td>
<td>210</td>
<td>189</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>188</td>
<td>203</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>196</td>
<td>191</td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>241</td>
<td>252</td>
<td></td>
</tr>
<tr>
<td>Pack-years smoking&lt;sup&gt;b&lt;/sup&gt;</td>
<td>802</td>
<td>799</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Moderate or strenuous activity (min/wk)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>824</td>
<td>827</td>
<td>0.05</td>
</tr>
<tr>
<td>Use of postmenopausal hormone therapy&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Never</td>
<td>415</td>
<td>346</td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>138</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>281</td>
<td>354</td>
<td>0.17</td>
</tr>
<tr>
<td>Family history of CRC (yes)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>167</td>
<td>143</td>
<td></td>
</tr>
<tr>
<td>History of colonoscopy or sigmoidoscopy (yes)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>431</td>
<td>500</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>History of colon polyp removal (yes)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>102</td>
<td>90</td>
<td>0.03</td>
</tr>
<tr>
<td>Plasma choline (µmol/L)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>835</td>
<td>835</td>
<td>0.25</td>
</tr>
<tr>
<td>Plasma betaine (µmol/L)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>835</td>
<td>835</td>
<td>0.31</td>
</tr>
<tr>
<td>Plasma DMG (µmol/L)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>835</td>
<td>834</td>
<td>0.89</td>
</tr>
<tr>
<td>Plasma TMAO (µmol/L)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>835</td>
<td>835</td>
<td>0.005</td>
</tr>
</tbody>
</table>

<sup>a</sup> Data are presented as n, value.:<sup>b</sup> Mean ± SD. <sup>c</sup> Count, %.<sup>d</sup> Median (IQR).
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma betaine:choline ratio</td>
<td>2.9 ± 1.2</td>
<td>3.0 ± 1.3</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Plasma DMG:choline ratio</td>
<td>0.27 ± 0.12</td>
<td>0.28 ± 0.11</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Plasma DMG:betaine ratio</td>
<td>0.10 ± 0.05</td>
<td>0.10 ± 0.05</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Plasma folate (ng/mL)d</td>
<td>15.6 (8.9-25.3)</td>
<td>17.2 (9.9-27.1)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>RBC folate (ng/mL)d</td>
<td>564 (410-742)</td>
<td>591 (431-751)</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Plasma PLP (nmol/L)d</td>
<td>60 (39-101)</td>
<td>67 (44-113)</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Plasma vitamin B12 (pg/mL)d</td>
<td>477 (336-661)</td>
<td>505 (376-691)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Plasma homocysteine (µmol/L)d</td>
<td>8.1 (6.8-9.9)</td>
<td>7.7 (6.7-9.4)</td>
<td>0.002</td>
<td></td>
</tr>
</tbody>
</table>

* Differences between cases and controls were analyzed by t tests (normally distributed continuous variables); Wilcoxon tests (non-normally distributed continuous variables); and chi-square tests (categorical variables).

* Values are mean ± SD for normally distributed continuous variables.

* Values are percentage for categorical variables.

* Values are median (interquartile range) for non-normally distributed continuous variables.

* Black or African-American, Hispanic, Asian or Pacific Islander, American Indian or Alaskan Native, or missing.
Correlations among plasma concentrations of choline metabolites

Spearman correlation coefficients \((r)\) were computed to examine associations among plasma choline metabolites. There were statistically significant, but modest, positive associations of plasma choline with plasma betaine \((r = 0.22; P < 0.001)\), DMG\((r = 0.21; P < 0.001)\) and TMAO\((r = 0.18; P < 0.001)\). Plasma betaine was also positively correlated with plasma DMG \((r = 0.39; P < 0.001)\).

Associations between plasma choline metabolites and colorectal cancer risk

In multivariable-adjusted analyses, women in the highest (vs. lowest) choline quartile were at an estimated 2.4 times greater risk of rectal cancer \((P\) trend = 0.08; Table 2.2). Conversely, women in the highest (vs. lowest) betaine quartile were at 32% lower colorectal cancer risk overall [OR (95% CI)_{highest vs. lowest quartile} = 0.68 (0.47–0.99); \(P\) trend = 0.01], 36% lower risk of local/regional tumors [0.64 (0.42–0.99); \(P\) trend = 0.009], and 31% lower risk of proximal tumors [0.69 (0.43–1.10); \(P\) trend = 0.05; Table 2.3]. No association between DMG quartiles and colorectal cancer risk was observed (Supplementary Table S2.1).
Table 2.2. ORs (95% CIs) of CRC by quartile of plasma cholinea

<table>
<thead>
<tr>
<th>Quartiles of choline (µmol/L)</th>
<th>1 (≤7.9)</th>
<th>2 (&gt;7.9-9.2)</th>
<th>3 (&gt;9.2-10.6)</th>
<th>4 (&gt;10.6)</th>
<th>P-trendb</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>412</td>
<td>403</td>
<td>408</td>
<td>447</td>
<td></td>
</tr>
<tr>
<td>All participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.06 (0.80, 1.40)</td>
<td>0.96 (0.71, 1.29)</td>
<td>1.30 (0.97, 1.74)</td>
<td>0.09</td>
</tr>
<tr>
<td>Multivariablec</td>
<td>1</td>
<td>1.01 (0.74, 1.39)</td>
<td>0.95 (0.68, 1.31)</td>
<td>1.22 (0.88, 1.70)</td>
<td>0.26</td>
</tr>
<tr>
<td>By tumor site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.16 (0.80, 1.70)</td>
<td>1.11 (0.75, 1.62)</td>
<td>1.33 (0.91, 1.95)</td>
<td>0.17</td>
</tr>
<tr>
<td>Multivariablec</td>
<td>1</td>
<td>1.06 (0.68, 1.66)</td>
<td>1.07 (0.69, 1.65)</td>
<td>1.21 (0.78, 1.87)</td>
<td>0.39</td>
</tr>
<tr>
<td>Distal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.02 (0.58, 1.82)</td>
<td>0.69 (0.34, 1.39)</td>
<td>1.12 (0.61, 2.05)</td>
<td>0.73</td>
</tr>
<tr>
<td>Multivariablec</td>
<td>1</td>
<td>0.92 (0.48, 1.77)</td>
<td>0.68 (0.31, 1.49)</td>
<td>1.07 (0.51, 2.23)</td>
<td>0.91</td>
</tr>
<tr>
<td>Rectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.08 (0.56, 2.08)</td>
<td>1.00 (0.51, 1.95)</td>
<td>1.79 (0.88, 3.64)</td>
<td>0.13</td>
</tr>
<tr>
<td>Multivariablec</td>
<td>1</td>
<td>1.38 (0.59, 3.22)</td>
<td>1.37 (0.56, 3.34)</td>
<td>2.44 (0.93, 6.40)</td>
<td>0.08</td>
</tr>
<tr>
<td>By stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local/regional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.11 (0.82, 1.51)</td>
<td>1.07 (0.78, 1.48)</td>
<td>1.33 (0.97, 1.81)</td>
<td>0.08</td>
</tr>
<tr>
<td>Multivariablec</td>
<td>1</td>
<td>1.01 (0.71, 1.44)</td>
<td>1.01 (0.70, 1.45)</td>
<td>1.23 (0.86, 1.76)</td>
<td>0.24</td>
</tr>
<tr>
<td>Metastatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.84 (0.37, 1.90)</td>
<td>0.41 (0.16, 1.04)</td>
<td>1.12 (0.46, 2.73)</td>
<td>0.82</td>
</tr>
<tr>
<td>Multivariablec</td>
<td>1</td>
<td>1.66 (0.56, 4.92)</td>
<td>0.55 (0.18, 1.73)</td>
<td>2.32 (0.69, 7.83)</td>
<td>0.30</td>
</tr>
</tbody>
</table>

aORs (95% CIs) of CRC were determined by conditional logistic regression.
bMedians for each quartile used in trend test: quartile 1 = 7.0 µmol/L, quartile 2 = 8.6 µmol/L, quartile 3 = 9.8 µmol/L, and quartile 4 = 11.8 µmol/L.
cMultivariable analyses were adjusted for age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, plasma vitamin B12, and plasma homocysteine.
Table 2.3. ORs (95% CIs) of CRC by quartile of plasma betaine

<table>
<thead>
<tr>
<th>Quartiles of betaine (µmol/L)</th>
<th>1 (≤18.8)</th>
<th>2 (&gt;18.8-26.6)</th>
<th>3 (≥26.6-34.0)</th>
<th>4 (≥34.0)</th>
<th>P-trend(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>413</td>
<td>464</td>
<td>417</td>
<td>376</td>
<td></td>
</tr>
</tbody>
</table>

**All participants**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.32 (1.01, 1.73)</td>
<td>1.02 (0.77, 1.36)</td>
<td>0.93 (0.70, 1.24)</td>
<td>0.29</td>
</tr>
<tr>
<td>Multivariable(^c)</td>
<td>1</td>
<td>1.03 (0.75, 1.43)</td>
<td>0.74 (0.52, 1.06)</td>
<td>0.68 (0.47, 0.99)</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**By tumor site**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proximal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.33 (0.95, 1.87)</td>
<td>1.07 (0.74, 1.54)</td>
<td>0.80 (0.55, 1.17)</td>
<td>0.16</td>
</tr>
<tr>
<td>Multivariable(^c)</td>
<td>1</td>
<td>1.26 (0.84, 1.89)</td>
<td>0.87 (0.55, 1.38)</td>
<td>0.69 (0.43, 1.10)</td>
<td>0.05</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.51 (0.81, 2.81)</td>
<td>1.25 (0.67, 2.36)</td>
<td>1.12 (0.58, 2.16)</td>
<td>0.95</td>
</tr>
<tr>
<td>Multivariable(^c)</td>
<td>1</td>
<td>0.89 (0.37, 2.11)</td>
<td>0.82 (0.33, 2.02)</td>
<td>0.63 (0.23, 1.73)</td>
<td>0.32</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rectal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.44 (0.74, 2.80)</td>
<td>0.65 (0.33, 1.27)</td>
<td>1.13 (0.60, 2.14)</td>
<td>0.71</td>
</tr>
<tr>
<td>Multivariable(^c)</td>
<td>1</td>
<td>1.02 (0.43, 2.42)</td>
<td>0.35 (0.13, 0.96)</td>
<td>0.61 (0.22, 1.70)</td>
<td>0.16</td>
</tr>
</tbody>
</table>

**By stage**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Local/regional</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.31 (0.98, 1.74)</td>
<td>0.91 (0.67, 1.23)</td>
<td>0.93 (0.67, 1.28)</td>
<td>0.23</td>
</tr>
<tr>
<td>Multivariable(^c)</td>
<td>1</td>
<td>1.01 (0.71, 1.44)</td>
<td>0.64 (0.43, 0.96)</td>
<td>0.64 (0.42, 0.99)</td>
<td>0.009</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metastatic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.34 (0.57, 3.15)</td>
<td>2.13 (0.87, 5.25)</td>
<td>0.91 (0.45, 1.85)</td>
<td>0.55</td>
</tr>
<tr>
<td>Multivariable(^c)</td>
<td>1</td>
<td>0.97 (0.33, 2.82)</td>
<td>1.91 (0.61, 5.95)</td>
<td>0.85 (0.31, 2.37)</td>
<td>0.70</td>
</tr>
</tbody>
</table>

\(^a\)ORs (95% CIs) of CRC were determined by conditional logistic regression.

\(^b\)Medians for each quartile used in trend test: quartile 1 = 14.4 µmol/L, quartile 2 = 22.8 µmol/L, quartile 3 = 29.9 µmol/L, and quartile 4 = 39.1 µmol/L.

\(^c\)Multivariable analyses were adjusted for age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, plasma vitamin B12, and plasma homocysteine.
Notably, after controlling for covariates, women in the highest (vs. lowest) quartile of the plasma betaine:choline ratio were at an estimated 44% lower colorectal cancer risk overall [0.56 (0.39–0.82); \( P \) trend = 0.004] as well as 34% lower risk of proximal tumors [0.66 (0.41–1.06); \( P \) trend = 0.049], 73% lower risk of rectal tumors [0.27 (0.10–0.78); \( P \) trend = 0.02], and 50% lower risk of local/regional tumors [0.27 (0.10–0.78); \( P \) trend = 0.001; Table 2.4]. The plasma DMG:choline ratio tended to be inversely associated with colorectal cancer risk overall [0.69 (0.48–0.98); \( P \) trend = 0.06; Supplementary Table S2.2]. The inverse association was statistically significant for local/regional tumors [0.62 (0.42–0.91); \( P \) trend = 0.04] and borderline significant for proximal tumors [0.57 (0.36–0.93); \( P \) trend = 0.07]. Last, the DMG:betaine ratio tended to be positively associated with rectal cancer risk [2.56 (0.98–6.64); \( P \) trend = 0.09; Supplementary Table S2.3].
Table 2.4. ORs (95% CIs) of CRC by quartile of plasma betaine:choline ratio

<table>
<thead>
<tr>
<th>Quartiles of betaine:choline ratio</th>
<th>1 (≤2.0)</th>
<th>2 (&gt;2.0-2.8)</th>
<th>3 (&gt;2.8-3.8)</th>
<th>4 (&gt;3.8)</th>
<th>P-trend&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>416</td>
<td>446</td>
<td>436</td>
<td>372</td>
<td></td>
</tr>
<tr>
<td>All participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.12 (0.85, 1.48)</td>
<td>1.08 (0.83, 1.41)</td>
<td>0.79 (0.59, 1.05)</td>
<td>0.08</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.83 (0.60, 1.15)</td>
<td>0.87 (0.62, 1.22)</td>
<td>0.56 (0.39, 0.82)</td>
<td>0.004</td>
</tr>
<tr>
<td>By tumor site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.26 (0.88, 1.79)</td>
<td>1.09 (0.77, 1.55)</td>
<td>0.74 (0.51, 1.09)</td>
<td>0.08</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>1.12 (0.73, 1.70)</td>
<td>0.98 (0.63, 1.53)</td>
<td>0.66 (0.41, 1.06)</td>
<td>0.049</td>
</tr>
<tr>
<td>Distal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.90 (0.48, 1.69)</td>
<td>1.07 (0.61, 1.87)</td>
<td>0.83 (0.43, 1.60)</td>
<td>0.76</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.53 (0.24, 1.18)</td>
<td>0.86 (0.40, 1.84)</td>
<td>0.45 (0.19, 1.10)</td>
<td>0.24</td>
</tr>
<tr>
<td>Rectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.06 (0.55, 2.06)</td>
<td>0.94 (0.51, 1.73)</td>
<td>0.75 (0.39, 1.41)</td>
<td>0.32</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.56 (0.22, 1.43)</td>
<td>0.45 (0.18, 1.13)</td>
<td>0.27 (0.10, 0.78)</td>
<td>0.02</td>
</tr>
<tr>
<td>By stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local/regional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.17 (0.86, 1.58)</td>
<td>1.04 (0.78, 1.38)</td>
<td>0.74 (0.54, 1.02)</td>
<td>0.04</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.88 (0.61, 1.27)</td>
<td>0.81 (0.56, 1.18)</td>
<td>0.50 (0.33, 0.76)</td>
<td>0.001</td>
</tr>
<tr>
<td>Metastatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.81 (0.38, 1.75)</td>
<td>1.05 (0.49, 2.24)</td>
<td>0.95 (0.43, 2.10)</td>
<td>0.95</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.55 (0.20, 1.54)</td>
<td>0.80 (0.27, 2.32)</td>
<td>0.79 (0.25, 2.50)</td>
<td>0.98</td>
</tr>
</tbody>
</table>

<sup>a</sup>ORs (95% CIs) of CRC were determined by conditional logistic regression.

<sup>b</sup>Medians for each quartile used in trend test: quartile 1 = 1.6, quartile 2 = 2.4, quartile 3 = 3.2 and quartile 4 = 4.4.

<sup>c</sup>Multivariable analyses were adjusted for age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, plasma vitamin B12, and plasma homocysteine.
Plasma TMAO, an oxidative derivative of choline produced by intestinal bacteria, was positively associated with colorectal cancer risk in age-adjusted analyses [1.78 (1.32–2.40); $P$ trend = 0.005; Table 2.5]. Women in the highest (vs. lowest) TMAO quartile were at approximately 1.9 times greater risk of proximal tumors ($P$ trend = 0.04), 2.3 times greater risk of rectal tumors ($P$ trend = 0.02), and 1.8 times greater risk of local/regional tumors ($P$ trend = 0.008). After controlling for covariates, the positive association remained strong and statistically significant for rectal cancer with approximately 3.4 times greater risk among women in the highest (vs. lowest) TMAO quartile ($P$ trend = 0.02). A borderline significant positive association was also observed for local/regional tumors with approximately 1.8 times greater risk in the highest (vs. lowest) TMAO quartile ($P$ trend = 0.08). Notably, although the linear trend across TMAO quartiles was not statistically significant, higher risk was observed from the second (vs. lowest) quartile of TMAO for colorectal cancer overall [1.90 (1.36–2.64)] and for proximal tumors [2.37 (1.52–3.70)]. Similarly, women in the second (vs. lowest) quartile of TMAO were at an estimated 1.9 times higher risk for local/regional tumors and 3.6 times higher risk for metastatic tumors, but this was not consistently observed in the other quartiles.
Table 2.5. ORs (95% CIs) of CRC by quartile of plasma TMAO

<table>
<thead>
<tr>
<th>Quartiles of TMAO (µmol/L)</th>
<th>1 (≤2.6)</th>
<th>2 (&gt;2.6-3.7)</th>
<th>3 (&gt;3.7-5.6)</th>
<th>4 (&gt;5.6)</th>
<th>P-trend&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>358</td>
<td>435</td>
<td>426</td>
<td>451</td>
<td></td>
</tr>
<tr>
<td>All participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.67 (1.25, 2.23)</td>
<td>1.55 (1.16, 2.07)</td>
<td>1.78 (1.32, 2.40)</td>
<td>0.005</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>1.90 (1.36, 2.64)</td>
<td>1.47 (1.06, 2.05)</td>
<td>1.65 (1.17, 2.34)</td>
<td>0.13</td>
</tr>
<tr>
<td>By tumor site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>2.06 (1.40, 3.03)</td>
<td>2.06 (1.39, 3.04)</td>
<td>1.93 (1.31, 2.83)</td>
<td>0.04</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>2.37 (1.52, 3.70)</td>
<td>1.92 (1.23, 3.00)</td>
<td>1.69 (1.09, 2.63)</td>
<td>0.42</td>
</tr>
<tr>
<td>Distal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.50 (0.77, 2.92)</td>
<td>1.20 (0.63, 2.27)</td>
<td>1.54 (0.78, 3.06)</td>
<td>0.41</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>1.96 (0.86, 4.48)</td>
<td>1.19 (0.56, 2.53)</td>
<td>1.69 (0.73, 3.90)</td>
<td>0.59</td>
</tr>
<tr>
<td>Rectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.03 (0.53, 1.98)</td>
<td>0.99 (0.52, 1.89)</td>
<td>2.26 (1.06, 4.79)</td>
<td>0.02</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>1.42 (0.62, 3.28)</td>
<td>1.20 (0.53, 2.72)</td>
<td>3.38 (1.25, 9.16)</td>
<td>0.02</td>
</tr>
<tr>
<td>By stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local/regional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.59 (1.16, 2.19)</td>
<td>1.56 (1.13, 2.14)</td>
<td>1.78 (1.28, 2.46)</td>
<td>0.008</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>1.90 (1.31, 2.74)</td>
<td>1.46 (1.00, 2.11)</td>
<td>1.78 (1.21, 2.60)</td>
<td>0.08</td>
</tr>
<tr>
<td>Metastatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>2.81 (1.23, 6.41)</td>
<td>1.61 (0.78, 3.32)</td>
<td>2.26 (0.96, 5.31)</td>
<td>0.17</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>3.63 (1.29, 10.23)</td>
<td>2.27 (0.86, 5.96)</td>
<td>2.09 (0.63, 6.97)</td>
<td>0.47</td>
</tr>
</tbody>
</table>

<sup>a</sup>ORs (95% CIs) of CRC were determined by conditional logistic regression.

<sup>b</sup>Medians for each quartile used in trend test: quartile 1 = 2.0 µmol/L, quartile 2 = 3.1 µmol/L, quartile 3 = 4.5 µmol/L, and quartile 4 = 8.1 µmol/L.

<sup>c</sup>Multivariable analyses were adjusted for age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, plasma vitamin B12, and plasma homocysteine.
**Associations of choline metabolites with colorectal cancer risk according to plasma B-vitamin concentrations**

To further explore whether B vitamins (folate, PLP, and vitamin B12) modified the associations between choline metabolites and colorectal cancer risk, we stratified into high/low plasma concentrations of B vitamins and assessed the interaction. After controlling for covariates, vitamin B12 status modified the association between plasma TMAO and colorectal cancer risk ($P$ interaction = 0.003; Table 2.6). Specifically, higher colorectal cancer risk was observed with higher TMAO quartiles among women with low plasma vitamin B12 (i.e., \( \leq 505 \) pg/mL; $P$ trend = 0.001), but not among those with high B12 levels. Other than this finding, no effect modifications by B vitamins were observed on the associations of plasma choline metabolites and their ratios with colorectal cancer risk (data not shown).

**Associations of choline metabolites with colorectal cancer risk according to FA fortification period**

We next explored the possible effect modification by FA fortification. The association of plasma choline, DMG, TMAO, and the ratios of choline metabolites with colorectal cancer risk did not differ by fortification periods (data not shown). However, after controlling for covariates, plasma betaine tended to interact with FA fortification period in association with colorectal cancer risk ($P$ interaction = 0.08; Table 2.7). Specifically, lower colorectal cancer risk was observed with higher plasma betaine during the pre- ($P$ trend = 0.02) and peri- ($P$ trend = 0.02) fortification periods, but not during the postfortification period.
Table 2.6. ORs (95% CIs) of CRC associated with quartiles of plasma TMAO by vitamin B12 status

<table>
<thead>
<tr>
<th>Vitamin B12 status</th>
<th>Quartiles of TMAO (µmol/L)</th>
<th>1 (≤ 2.6)</th>
<th>2 (&gt;2.6-3.7)</th>
<th>3 (&gt;3.7-5.6)</th>
<th>4 (&gt;5.6)</th>
<th>P-interactionc</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Age-adjusted</td>
<td>Multivariabled</td>
<td>Age-adjusted</td>
<td>Multivariabled</td>
<td>Age-adjusted</td>
</tr>
<tr>
<td>Low B12 (≤505 pg/mL)</td>
<td></td>
<td>77</td>
<td>107</td>
<td>122</td>
<td>153</td>
<td>0.0007</td>
</tr>
<tr>
<td>no. of cases</td>
<td></td>
<td>1</td>
<td>1.74 (1.17, 2.58)</td>
<td>2.01 (1.35, 2.98)</td>
<td>2.49 (1.68, 3.67)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2.00 (1.30, 3.06)</td>
<td>2.06 (1.34, 3.17)</td>
<td>2.44 (1.59, 3.75)</td>
<td></td>
</tr>
<tr>
<td>High B12 (&gt;505 pg/mL)</td>
<td></td>
<td>71</td>
<td>122</td>
<td>95</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>no. of cases</td>
<td></td>
<td>1</td>
<td>1.45 (0.97, 2.18)</td>
<td>1.11 (0.73, 1.69)</td>
<td>1.00 (0.66, 1.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>1.49 (0.96, 2.32)</td>
<td>0.98 (0.63, 1.55)</td>
<td>0.92 (0.58, 1.47)</td>
<td></td>
</tr>
</tbody>
</table>

aORs (95% CIs) of CRC were determined by unconditional logistic regression due to case-control matching being broken in these subset analyses. Models were additionally adjusted for ethnicity and time to diagnosis.

bMedians for each quartile: quartile 1 = 2.0 µmol/L, quartile 2 = 3.1 µmol/L, quartile 3 = 4.5 µmol/L, and quartile 4 = 8.1 µmol/L.

cP value for test of interaction between TMAO (as an ordinal variable) and plasma B-vitamin status.

dMultivariable analyses were adjusted for days to CRC diagnosis, ethnicity, age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, and plasma homocysteine.
Table 2.7. ORs (95% CIs) of CRC associated with quartiles of plasma betaine by FA fortification periodsa

<table>
<thead>
<tr>
<th>Fortification period</th>
<th>Quartiles of betaine (µmol/L)b</th>
<th>1(≤18.8)</th>
<th>2 (&gt;18.8-26.6)</th>
<th>3 (&gt;26.6-34.0)</th>
<th>4 (&gt;34.0)</th>
<th>P-interactionc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-fortification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>no. of cases</td>
<td>50</td>
<td>65</td>
<td>49</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age-adjusted</td>
<td>1</td>
<td>1.45 (0.84, 2.51)</td>
<td>0.85 (0.49, 1.48)</td>
<td>0.73 (0.41, 1.29)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multivariabled</td>
<td>1</td>
<td>1.06 (0.55, 2.01)</td>
<td>0.65 (0.32, 1.31)</td>
<td>0.46 (0.22, 0.98)</td>
<td></td>
</tr>
<tr>
<td>Peri-fortification</td>
<td>no. of cases</td>
<td>107</td>
<td>147</td>
<td>116</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>no. of cases</td>
<td>1</td>
<td>1.43 (0.99, 2.07)</td>
<td>0.98 (0.67, 1.42)</td>
<td>0.78 (0.53, 1.15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multivariabled</td>
<td>1</td>
<td>1.10 (0.72, 1.67)</td>
<td>0.74 (0.47, 1.15)</td>
<td>0.64 (0.39, 1.04)</td>
<td></td>
</tr>
<tr>
<td>Post-fortification</td>
<td>no. of cases</td>
<td>44</td>
<td>48</td>
<td>38</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>no. of cases</td>
<td>1</td>
<td>1.09 (0.62, 1.92)</td>
<td>1.39 (0.74, 2.60)</td>
<td>1.58 (0.85, 2.91)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multivariabled</td>
<td>1</td>
<td>0.88 (0.46, 1.69)</td>
<td>0.87 (0.41, 1.86)</td>
<td>0.97 (0.45, 2.06)</td>
<td></td>
</tr>
</tbody>
</table>

aORs (95% CIs) of CRC were determined by unconditional logistic regression due to case-control matching being broken in these subset analyses. Models were additionally adjusted for ethnicity and time to diagnosis.
bMedians for each quartile: quartile 1 = 14.4 µmol/L, quartile 2 = 22.8 µmol/L, quartile 3 = 29.9 µmol/L, and quartile 4 = 39.1 µmol/L.
cP value for test of interaction between betaine (as an ordinal variable) and FA fortification periods.
dMultivariable analyses were adjusted for days to CRC diagnosis, ethnicity, age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, plasma vitamin B12, and plasma homocysteine.
Discussion

To the best of our knowledge, this is the first study to assess associations between plasma biomarkers of choline metabolism and colorectal cancer risk among postmenopausal women in the United States. The following main findings emerged: (i) plasma choline (modest positive) and betaine (inverse) were divergently associated with colorectal cancer risk; (ii) the plasma betaine:choline ratio was more strongly associated with colorectal cancer risk than was either metabolite alone; and (iii) higher plasma TMAO concentrations were associated with higher risk of colorectal cancer especially among women with low plasma vitamin B12.

The divergent associations of plasma choline and betaine with colorectal cancer risk are unexpected given that betaine is derived from choline and increases in response to a higher choline intake (24). Thus, the divergent associations may arise from the disease process itself, which could alter choline metabolism before diagnosis (29, 30). For example, postmenopausal women harboring undiagnosed, precancerous lesions may have a higher demand for choline due to its greater use for membrane biosynthesis by abnormally dividing cells (31, 32). This in turn may upregulate de novo choline production through the hepatic phosphatidylethanolamine N-methyltransferase (PEMT) pathway. Enhanced hepatic PEMT activity would be expected to elevate choline, a product of the PEMT reaction, while depleting betaine, a source of methyl groups for the PEMT reaction. This metabolic scenario is observed during pregnancy (33), which like cancer is a state of rapidly dividing cells and exhibits several of the same molecular characteristics (34). However, unlike pregnancy where providing substrate for the PEMT reaction may beneficially influence fetal growth and development, betaine supplementation for the purposes of colorectal cancer reduction among postmenopausal women appears unwise because the prevalence of colonic neoplasia increases with age (35) and extra betaine may
accelerate tumor progression.

The divergent associations of plasma choline and betaine with colorectal cancer risk observed in our study cohort differ from findings of a recent case–control study nested within the EPIC cohort, where, in the subgroup analyses of women, plasma choline (but not plasma betaine) was inversely associated with colorectal cancer risk (13). One major difference between the study cohorts that could explain these discordant findings is folate status. Specifically, median plasma folate concentrations were approximately 3.5 times higher in the WHI (vs. EPIC) cohort. Other contributing factors may include age of participants, follow-up period, blood sample collection (fasting vs. nonfasting), use of different cutpoints for categories of choline metabolites, and the status of other nutrients involved in one-carbon metabolism.

In the present study, the plasma betaine:choline ratio was more strongly associated with colorectal cancer risk than either metabolite alone. After adjusting for potential confounders, women in the highest (vs. lowest) betaine:choline quartile were at 44% lower colorectal cancer risk overall, 34% lower proximal tumors, 50% lower local/regional tumors, and 73% lower rectal tumors. The association between the betaine:choline ratio and colorectal cancer risk did not appear to differ according to B-vitamin status or FA fortification period. In contrast, FA exposure appeared to modify the association between plasma betaine and colorectal cancer risk with an inverse association observed in the preand perifortification periods, but not in the postfortification period. As such, the association between plasma betaine and colorectal cancer risk appears to be dependent on folate availability and may be more evident when folate availability is low (i.e., before FA fortification). Overall, these data support the utility of the plasma betaine:choline ratio as a potential biomarker for excess risk of colorectal cancer in postmenopausal women.
In humans, choline can undergo catabolism by anaerobic intestinal bacteria to produce TMA, which is further converted to TMAO by the hepatic enzyme FMO (14, 15). Similarly, L-carnitine also serves as a precursor of TMAO through a gut microbiota–dependent metabolism (i.e., choline/carnitine → gut microbiota → TMA/TMAO; refs. 36, 37). This metabolic pathway mediated by intestinal microbiota has been linked to several diseases (37–41), suggesting the potential role of gut-microbial metabolism and their metabolic products in carcinogenesis among humans. The present study, for the first time to our knowledge, examined an association between circulating concentrations of TMAO and colorectal cancer risk. We found that women in the highest (vs. lowest) TMAO quartile had an approximately 3.4 times greater risk of rectal cancer. Although no statistically significant linear trend was observed, increased risk was also detected from the second quartile of TMAO with 1.9 times greater risk for colorectal cancer overall and for local/regional tumors, approximately 2.4 times greater risk for proximal tumors, and approximately 3.6 times greater risk for metastatic tumors. These findings collectively suggest that plasma TMAO may serve as a potential predictor of increased colorectal cancer risk.

Alterations in the intestinal microbiota may predispose to the development and progression of colorectal cancer through affecting multiple processes, including colonic epithelial cell proliferation, immune system, and chronic inflammation (16, 18). For example, compared with healthy individuals, increased number and diversity as well as the decreased stability of a colonic bacterial group, Clostridium, have been characterized in patients with colorectal cancer (16, 42). Indeed, Clostridium is also suggested to play a role in the conversion of choline (41, 43) and carnitine (37, 44) to TMA, thereby contributing to TMAO production. Thus, it is possible that the positive association between plasma TMAO and colorectal cancer risk may arise from abnormal changes in particular colonic bacteria, which could occur early in
disease development before diagnosis. Given that TMAO is a gut bacteria–derived metabolite, it may also represent evidence for an etiologic correlation between intestinal microbiota and colorectal cancer and could potentially serve as a novel biomarker of colorectal cancer risk.

Notably, the association between plasma TMAO and colorectal cancer risk appeared to be modified by vitamin B12 status. Specifically, the risk of colorectal cancer increased across increasing TMAO quartiles in the low B12 group, but not in the high B12 group. These data suggest that postmenopausal women with higher TMAO and lower vitamin B12 may be more susceptible to developing colorectal cancer. Certain groups of intestinal bacteria can synthesize (45, 46) and consume (47, 48) vitamin B12, which may affect the vitamin B12 requirement/status of the host. Indeed, overgrowth of intestinal bacteria that take up vitamin B12 has been implicated in B12 malabsorption (47–50). In human intestine, overgrowth of a specific bacterial group can also block colonization of other bacterial groups (16), yielding an imbalance between their metabolic production and consumption. Therefore, elevated colorectal cancer risk among women with high TMAO and low vitamin B12 may in part be associated with the disturbances in colonic bacterial populations. Additional studies are required to confirm these findings, and potential biologic mechanisms need further elucidation.

Key strengths of the present study include: (i) the prospective design; (ii) the large sample size, which allowed for stratified analyses by tumor site/stage as well as by B vitamins and FA fortification periods; and (iii) assessment of choline metabolite ratios (especially betaine:choline ratio), which provided more robust colorectal cancer risk estimates. Several limitations should also be noted: (i) although we attempted to control confounding, there is a potential for residual confounding by factors that were either not collected in the WHI-OS or not measured with sufficient precision; (ii) although the concentrations of plasma choline and its
metabolites are stable through time in healthy women (24), single measures of these metabolites may not fully reflect long-term associations with colorectal cancer risk; and (iii) although baseline hysterectomy status was used as a matching factor based on the evidence that female sex hormones (e.g., estrogen) are associated with colorectal cancer risk (51–53), it may not comprehensively account for estrogen status. However, this would not be expected to have an influence on the results, as the analyses were adjusted for the use of postmenopausal hormone therapy (which would more comprehensively account for estrogen status).

In conclusion, the results of this study indicate that alterations in choline metabolism, which may arise early in disease development, associate with higher risk of colorectal cancer in postmenopausal women. Our data also indicate that the plasma betaine:choline ratio may be a potential indicator of colorectal cancer risk, which, if confirmed, could have clinical implications for colorectal cancer screening. This study also provides new evidence that plasma TMAO, an oxidative derivative of choline produced by intestinal bacteria, may serve as a potential biomarker for increased risk of colorectal cancer especially among those with low plasma vitamin B12 concentrations. Although further investigations are needed to delineate the underlying mechanisms, these novel findings may advance understanding of an etiologic correlation between intestinal bacteria and colorectal cancer pathogenesis.
REFERENCES


26. Gilfix BM, Blank DW, Rosenblatt DS. Novel reductant for determination of total plasma


53. Grodstein F, Newcomb PA, Stampfer MJ. Postmenopausal hormone therapy and the risk
## Supplementary Table S2.1. ORs (95% CIs) of CRC by quartile of plasma DMG<sup>a</sup>

<table>
<thead>
<tr>
<th></th>
<th>Quartiles of DMG (µmol/L)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (≤1.9)</td>
<td>2 (&gt;1.9-2.4)</td>
<td>3 (&gt;2.4-2.9)</td>
<td>4 (&gt;2.9)</td>
<td>P-trend&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>427</td>
<td>424</td>
<td>396</td>
<td>422</td>
<td></td>
</tr>
<tr>
<td>All participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.99 (0.74, 1.31)</td>
<td>0.98 (0.72, 1.32)</td>
<td>1.03 (0.77, 1.40)</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.85 (0.61, 1.18)</td>
<td>0.79 (0.55, 1.11)</td>
<td>0.75 (0.53, 1.07)</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>By tumor site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.90 (0.62, 1.32)</td>
<td>0.93 (0.62, 1.39)</td>
<td>0.88 (0.60, 1.31)</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.75 (0.49, 1.17)</td>
<td>0.71 (0.44, 1.13)</td>
<td>0.67 (0.42, 1.06)</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Distal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.77 (0.42, 1.39)</td>
<td>0.99 (0.51, 1.90)</td>
<td>0.86 (0.45, 1.65)</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.53 (0.25, 1.12)</td>
<td>1.04 (0.47, 2.32)</td>
<td>0.51 (0.22, 1.17)</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Rectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.63 (0.81, 3.28)</td>
<td>1.10 (0.57, 2.12)</td>
<td>2.22 (1.06, 4.69)</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>2.17 (0.84, 5.62)</td>
<td>1.23 (0.51, 2.99)</td>
<td>2.46 (0.90, 6.77)</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>By stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local/regional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.96 (0.70, 1.32)</td>
<td>1.01 (0.73, 1.41)</td>
<td>1.03 (0.74, 1.43)</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.87 (0.60, 1.25)</td>
<td>0.84 (0.57, 1.24)</td>
<td>0.77 (0.52, 1.14)</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Metastatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.30 (0.59, 2.86)</td>
<td>0.86 (0.39, 1.88)</td>
<td>1.01 (0.44, 2.32)</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.96 (0.38, 2.43)</td>
<td>0.71 (0.26, 1.93)</td>
<td>0.89 (0.29, 2.73)</td>
<td>0.78</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>ORs (95% CIs) of CRC were determined by conditional logistic regression.

<sup>b</sup>Medians for each quartile used in trend test: quartile 1 = 1.6 µmol/L, quartile 2 = 2.1 µmol/L, quartile 3 = 2.6 µmol/L, and quartile 4 = 3.4 µmol/L.

<sup>c</sup>Multivariable analyses were adjusted for age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, plasma vitamin B12, and plasma homocysteine.
### Supplementary Table S2.2. ORs (95% CIs) of CRC by quartile of plasma DMG:choline ratio

<table>
<thead>
<tr>
<th>Quartiles of DMG:choline ratio</th>
<th>1 (≤0.20)</th>
<th>2 (&gt;0.20-0.25)</th>
<th>3 (&gt;0.25-0.32)</th>
<th>4 (&gt;0.32)</th>
<th>P-trend&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>431</td>
<td>403</td>
<td>433</td>
<td>402</td>
<td></td>
</tr>
<tr>
<td>All participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.86 (0.65, 1.14)</td>
<td>0.97 (0.73, 1.30)</td>
<td>0.83 (0.61, 1.14)</td>
<td>0.36</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.82 (0.60, 1.13)</td>
<td>0.93 (0.66, 1.30)</td>
<td>0.69 (0.48, 0.98)</td>
<td>0.06</td>
</tr>
<tr>
<td>By tumor site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.69 (0.47, 1.01)</td>
<td>0.80 (0.54, 1.18)</td>
<td>0.65 (0.43, 0.98)</td>
<td>0.10</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.64 (0.42, 0.99)</td>
<td>0.79 (0.50, 1.24)</td>
<td>0.57 (0.36, 0.93)</td>
<td>0.07</td>
</tr>
<tr>
<td>Distal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.93 (1.02, 3.67)</td>
<td>1.18 (0.65, 2.14)</td>
<td>1.39 (0.70, 2.78)</td>
<td>0.54</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>2.13 (1.02, 4.45)</td>
<td>1.15 (0.57, 2.33)</td>
<td>1.04 (0.47, 2.30)</td>
<td>0.75</td>
</tr>
<tr>
<td>Rectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.61 (0.33, 1.13)</td>
<td>1.04 (0.52, 2.07)</td>
<td>0.86 (0.40, 1.87)</td>
<td>0.96</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.62 (0.28, 1.37)</td>
<td>1.20 (0.50, 2.90)</td>
<td>0.98 (0.38, 2.56)</td>
<td>0.86</td>
</tr>
<tr>
<td>By stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local/regional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>0.78 (0.58, 1.06)</td>
<td>0.87 (0.64, 1.19)</td>
<td>0.75 (0.53, 1.05)</td>
<td>0.17</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.70 (0.49, 0.99)</td>
<td>0.81 (0.56, 1.16)</td>
<td>0.62 (0.42, 0.91)</td>
<td>0.04</td>
</tr>
<tr>
<td>Metastatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-adjusted</td>
<td>1</td>
<td>1.74 (0.75, 4.02)</td>
<td>2.54 (0.98, 6.59)</td>
<td>2.13 (0.74, 6.15)</td>
<td>0.21</td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>2.57 (0.85, 7.84)</td>
<td>2.96 (0.84, 10.38)</td>
<td>1.96 (0.46, 8.33)</td>
<td>0.50</td>
</tr>
</tbody>
</table>

<sup>a</sup>ORs (95% CIs) of CRC were determined by conditional logistic regression.

<sup>b</sup>Medians for each quartile used in trend test: quartile 1 = 0.17, quartile 2 = 0.23, quartile 3 = 0.28 and quartile 4 = 0.39.

<sup>c</sup>Multivariable analyses were adjusted for age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, plasma vitamin B12, and plasma homocysteine.
**Supplementary Table S2.3.** ORs (95% CIs) of CRC by quartile of plasma DMG:betaine ratio

<table>
<thead>
<tr>
<th>Quartiles of DMG:betaine ratio</th>
<th>1 (≤0.07)</th>
<th>2 (0.07-0.09)</th>
<th>3 (0.09-0.12)</th>
<th>4 (&gt;0.12)</th>
<th>P-trend&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>403</td>
<td>430</td>
<td>402</td>
<td>433</td>
<td></td>
</tr>
</tbody>
</table>

**All participants**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-adjusted</td>
<td>1.10 (0.84, 1.45)</td>
<td>0.96 (0.72, 1.28)</td>
<td>1.13 (0.85, 1.50)</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.07 (0.77, 1.47)</td>
<td>0.89 (0.64, 1.25)</td>
<td>1.08 (0.76, 1.55)</td>
<td>0.78</td>
<td></td>
</tr>
</tbody>
</table>

**By tumor site**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximal</td>
<td>1.08 (0.75, 1.54)</td>
<td>0.87 (0.60, 1.27)</td>
<td>1.04 (0.72, 1.51)</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.04 (0.68, 1.59)</td>
<td>0.90 (0.58, 1.38)</td>
<td>0.94 (0.59, 1.48)</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Distal</td>
<td>0.85 (0.45, 1.58)</td>
<td>1.02 (0.52, 1.98)</td>
<td>1.08 (0.57, 2.04)</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.77 (0.37, 1.61)</td>
<td>0.78 (0.34, 1.82)</td>
<td>1.06 (0.45, 2.51)</td>
<td>0.73</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectal</td>
<td>1.38 (0.73, 2.63)</td>
<td>1.23 (0.65, 2.31)</td>
<td>1.53 (0.79, 2.94)</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.01 (0.85, 4.73)</td>
<td>1.32 (0.56, 3.11)</td>
<td>2.56 (0.98, 6.64)</td>
<td>0.09</td>
<td></td>
</tr>
</tbody>
</table>

**By stage**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Local/regional</td>
<td>0.99 (0.73, 1.34)</td>
<td>0.94 (0.69, 1.29)</td>
<td>1.06 (0.78, 1.45)</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.95 (0.66, 1.36)</td>
<td>0.88 (0.61, 1.27)</td>
<td>1.07 (0.72, 1.58)</td>
<td>0.69</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Metastatic</td>
<td>1.70 (0.84, 3.42)</td>
<td>0.96 (0.45, 2.05)</td>
<td>1.78 (0.82, 3.91)</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Multivariable&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.55 (0.63, 3.82)</td>
<td>0.79 (0.26, 2.39)</td>
<td>1.48 (0.49, 4.48)</td>
<td>0.66</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>ORs (95% CIs) of CRC were determined by conditional logistic regression.

<sup>b</sup>Medians for each quartile used in trend test: quartile 1 = 0.06, quartile 2 = 0.08, quartile 3 = 0.10 and quartile 4 = 0.15.

<sup>c</sup>Multivariable analyses were adjusted for age, baseline BMI, pack-years of smoking, moderate or strenuous physical activity (min/wk), use of postmenopausal-hormone-therapy, history of colonoscopy, RBC folate, plasma PLP, plasma vitamin B12, and plasma homocysteine.
CHAPTER 3

Vitamin B-12 status differs among pregnant, lactating, and control women with equivalent nutrient intakes*

ABSTRACT

**Background:** Limited data are available from controlled studies on biomarkers of maternal vitamin B-12 status.

**Objective:** We sought to quantify the effects of pregnancy and lactation on the vitamin B-12 status response to a known and highly controlled vitamin B-12 intake.

**Methods:** As part of a 10–12 wk feeding trial, pregnant (26–29 wk gestation; n = 26), lactating (5 wk postpartum; n = 28), and control (nonpregnant, nonlactating; n = 21) women consumed vitamin B-12 amounts of ~8.6 µg/d [mixed diet (~6 µg/d) plus a prenatal multivitamin supplement (2.6 µg/d)]. Serum vitamin B-12, holotranscobalamin (bioactive form of vitamin B-12), methylmalonic acid (MMA), and homocysteine were measured at baseline and study-end.

**Results:** All participants achieved adequate vitamin B-12 status in response to the study dose. Compared with control women, pregnant women had lower serum vitamin B-12 (221%; \( P = 0.02 \)) at study-end, whereas lactating women had higher (\( P = 0.04 \)) serum vitamin B-12 throughout the study (+26% at study-end). Consumption of the study vitamin B-12 dose increased serum holotranscobalamin in all reproductive groups (+16–42%; \( P \leq 0.009 \)). At study-end, pregnant (vs. control) women had a higher holotranscobalamin–to–vitamin B-12 ratio (\( P = 0.04 \)) with ~30% (vs. 20%) of total vitamin B-12 in the bioactive form. Serum MMA increased during pregnancy (+50%; \( P < 0.001 \)) but did not differ by reproductive state at study-end. Serum homocysteine increased in pregnant women (+15%; \( P = 0.009 \)) but decreased in control and lactating women (-16–17%; \( P < 0.001 \)). Despite these changes, pregnant women had ~20% lower serum homocysteine than the other 2 groups at study-end (\( P \leq 0.02 \)).

**Conclusion:** Pregnancy and lactation alter vitamin B-12 status in a manner consistent with enhanced vitamin B-12 supply to the child. Consumption of the study vitamin B-12 dose (~3
times the RDA) increased the bioactive form of vitamin B-12, suggesting that women in these reproductive states may benefit from vitamin B-12 intakes exceeding current recommendations. This trial was registered at clinicaltrials.gov as NCT01127022.
INTRODUCTION

Vitamin B-12, a water-soluble micronutrient, is essential for hematologic and neurologic processes. It serves as a cofactor in the remethylation of homocysteine to methionine and in the conversion of L-methylmalonyl-CoA to succinyl-CoA. The current RDA of vitamin B-12 is 2.4 µg/d for US adults with upward adjustments to 2.6 and 2.8 µg/d during pregnancy and lactation, respectively (1). However, previous studies conducted in healthy adults have suggested that vitamin B-12 intakes greater than the current recommendations may be required to ensure optimal vitamin B-12 status (2–5).

Maternal vitamin B-12 deficiency has been associated with the increased risk of adverse pregnancy outcomes (e.g., neural tube defects, preterm delivery, and intrauterine growth retardation) (6–9), indicating the importance of sufficient vitamin B-12 intake/status during pregnancy for optimal fetal development and growth. Maternal serum vitamin B-12 concentrations gradually decline throughout normal pregnancy with the lowest concentration reached in late gestation (10–13). Although a few studies have estimated dietary and/or supplemental vitamin B-12 intakes during this reproductive state (11, 13), to the best of our knowledge, no studies have assessed vitamin B-12 status of pregnant women under controlled feeding conditions.

Maternal vitamin B-12 intake/status during lactation may influence the vitamin B-12 concentration of her milk and possibly the vitamin B-12 status of breastfed infants (14). However, it has also been suggested that even high doses of maternal vitamin B-12 supplementation during lactation may not yield a significant increase in milk vitamin B-12 concentrations, thus not affecting infant vitamin B-12 status (15). Infantile vitamin B-12 deficiency can lead to neurologic impairments including irritability, apathy, and developmental
regression (16–18). Although a few studies have examined longitudinal changes in maternal vitamin B-12 biomarkers during lactation (14, 19–21), less is known about the impact of this reproductive state (compared to the nonlactating state) on vitamin B-12 status and requirement.

The objective of the present study was to quantify the effects of pregnancy and lactation on vitamin B-12 status biomarkers under controlled feeding conditions. To the best of our knowledge, this was the first study to assess and compare vitamin B-12 status response among pregnant, lactating, and control (nonpregnant, nonlactating) women with equivalent vitamin B-12 intakes. We also assessed breast milk vitamin B-12 concentration in response to a known vitamin B-12 dose and its associations with maternal vitamin B-12 biomarkers.
METHODS

Participants and study design

The present study was an extension of a 10–12 wk controlled feeding trial (22, 23) in which healthy pregnant (26–29 wk gestation; n = 26), lactating (5 wk postpartum; n = 28), and control (nonpregnant, nonlactating; n = 21) women, aged 21–40 y, were randomly assigned to choline intakes of either 480 or 930 mg/d (380-mg choline/d from diet; 100 or 550 mg/d from supplemental choline chloride). The study participants were recruited from Ithaca, New York, and surrounding areas between January 2009 and October 2010 as previously described (22, 23). During the screening stage, interested women completed a questionnaire on their age, ethnicity/race, prepregnancy weight (for pregnant and lactating women), multivitamin supplement use, medication use, health history, education, work status, physical activity, and tobacco/drug/alcohol use. Inclusion criteria for study enrollment were the following: 1) general healthy status as assessed by the questionnaire, blood chemistry profile, and complete blood count; 2) normal kidney and liver function; and 3) willingness to comply with the study protocol (i.e., agreement to eat >3 meals/wk at the onsite location and only consume food and beverages provided by the study). An additional inclusion criterion for lactating women was the intention to exclusively breastfeed for the duration of the study. Women were excluded if they were taking prescription medications known to affect liver function; if they used tobacco, recreational drugs, or alcohol (for pregnant and lactating women); or if they had a history of chronic disease. Additional exclusion criteria for pregnant women were nonsingleton pregnancy and pregnancy-associated complications (e.g., preeclampsia or gestational diabetes) (22, 23). The study protocol was reviewed and approved by the Institutional Review Board for Human Study Participant Use at Cornell University and at Cayuga Medical Center where pregnant participants delivered their
Study diet and supplements

Throughout the study period of 10–12 wk, all participants consumed equivalent vitamin B-12 amounts of ~8.6 µg/d, which is ~3–3.5 times the RDA of vitamin B-12 for adult females (2.4 µg/d), pregnant women (2.6 µg/d), and lactating women (2.8 µg/d) (1). The sources of vitamin B-12 intake were the study diet containing ~6-mg vitamin B-12/d and a daily over-the-counter prenatal multivitamin supplement (Pregnancy Plus; Fairhaven Health LLC) labeled to contain 2.6-µg vitamin B-12. The vitamin B-12 content of food and beverage items in the 7-d rotational study menu was calculated on the basis of the USDA National Nutrient Database for Standard Reference (24) (Supplemental Table S3.1).

As previously described (22, 23), the study meal provided ~2000 kcal/d and was prepared by study personnel in the Human Metabolic Research Unit at Cornell University. The caloric intake could be modified by participants with addition or subtraction of the following food items that did not contain vitamin B-12: unenriched white rice, snack items (i.e., chips, popcorn, rice cake, and apple), and beverages (i.e., soda, lemonade, and apple juice). In addition, all participants consumed a 200-mg DHA supplement on a daily basis (Neuromins; Nature’s Way Products) and a thrice-weekly potassium and magnesium supplement (General Nutrition Corp) to achieve nutrient intake recommendations (25, 26). Under the supervision of study personnel in the Human Metabolic Research Unit, pregnant and control women consumed at least one meal (and supplements) per day (Monday through Friday), whereas lactating women consumed one meal (and supplements) every other day. All other food, beverages, and supplements were provided as takeaways. To enhance study compliance, all participants completed a daily
checklist indicating that they received and consumed all menu items and supplements. In addition, study personnel had near-daily contact with participants throughout the study period to maintain positive rapport and monitor compliance.

**Sample collection and analytic measurements**

Blood samples were drawn at study baseline (week 0) and study-end (week 12 for pregnant and control participants; week 10 for lactating participants) after 10 h of fasting. The plasma and serum were separated and stored at -80°C until analysis as previously described (22, 23). Twenty-four-hour urine samples were also collected at baseline and study-end and stored at -80°C. For each collection, the total volume of urine was recorded. Lactating women were fasted (10 h) for milk sample collection, which occurred on the same day as the corresponding week’s blood draw and consisted of the full expression of one breast 2 h after the first feed of the day. Upon collection, samples were processed and stored at -80°C as previously described (23).

Serum vitamin B-12 was measured by automated chemiluminescence immunoassay with use of the Immulite 2000 (Siemens Medical Solutions Diagnostics); serum holotranscobalamin (bioactive form of vitamin B-12) was measured by the Axis-Shield Active-B12 EIA (Axis-Shield Diagnostics); and serum and urinary homocysteine as well as serum methylmalonic acid (MMA), functional indicators of vitamin B-12 status that increase upon intracellular depletion of vitamin B-12, were quantified by GC-MS (27, 28). Breast milk vitamin B-12 concentration was determined by a recently published competitive protein binding immunoassay with use of the Immulite 1000 (Siemens Medical Solutions Diagnostics) (29). This method yielded a small SD among the samples but a relatively low recovery rate (29), which may possibly underestimate breast milk vitamin B-12 concentrations. However, our own recovery experiments yielded a
mean recovery rate of ~105%, suggesting that breast milk vitamin B-12 was not underestimated in this study. Urinary creatinine was measured with use of the Dimension Xpand Clinical chemistry system (Siemens Healthcare Diagnostics) and used to normalize urinary homocysteine concentrations; and transcobalamin II (TCN2) C776G genotype (rs1801198), a common polymorphism known to influence vitamin B-12 status biomarkers (30, 31), was determined with use of a florescent TaqMan probe commercially available kit (Applied Biosystems). The mean interassay CVs of the internal laboratory control samples for each of the assays were as follows: serum vitamin B-12, 5%; serum holotranscobalamin, 4%; serum/urinary homocysteine, 3%; serum MMA, 7%; breast milk vitamin B-12, 4%; and urinary creatinine, 5%.

Statistical analysis

Baseline characteristics of the reproductive groups (i.e., pregnant, lactating, and control women) were compared with use of 1) one-factor ANOVA for normally distributed continuous variables; 2) nonparametric Kruskal-Wallis tests for non-normally distributed continuous variables; or 3) chi-square tests for categorical variables. When significant differences were observed, Bonferroni corrections were made for post hoc comparisons.

To test the effect of reproductive states on vitamin B-12 status biomarkers and to assess vitamin B-12 status response through time, linear mixed models were used. Data that deviated from the normal distribution or homogeneity of variance (i.e., all serum/urinary vitamin B-12 biomarkers and breast milk vitamin B-12) were natural log transformed and used in the models. In addition to serum vitamin B-12 and holotranscobalamin, their ratio (i.e., holotranscobalamin–to–vitamin B-12) was included in the models (after log transformation) as an additional vitamin B-12 status indicator because this ratio (vs. individual metabolite alone) is suggested to be...
reflect the tissue availability of circulating vitamin B-12 (32). Reproductive states (pregnant, lactating, or control), time (baseline or study-end), and the interaction of reproductive state and time (reproductive state x time) were entered as fixed factors, and participant identification was entered as a random factor. In the initial models, choline intake (480 or 930 mg/d) and potential confounding factors including age, ethnicity/race, baseline BMI (or pre-pregnant BMI for pregnant and lactating women), multivitamin supplement use before entering the study, and TCN2 C776G genotype were included; nonsignificant variables (\( P > 0.10 \)) were then removed in a stepwise manner until final models were derived. Specifically, BMI (\( P < 0.001 \)), multivitamin supplement use (\( P = 0.04 \)), choline intake (\( P = 0.04 \)), and TCN2 C776G genotype (\( P = 0.09 \)) were retained in the final models of serum vitamin B-12, serum holotranscobalamin, serum MMA, and urinary homocysteine (normalized to urinary creatinine), respectively. Ethnicity/race (\( P = 0.09 \)) and TCN2 C776G genotype (\( P = 0.10 \)) were also retained in the analysis of breast milk vitamin B-12. For other vitamin B-12 biomarkers (i.e., the holotranscobalamin-to–vitamin B-12 ratio and serum homocysteine), none of the covariates were included in final models. The analyses were then followed by post hoc contrasts (Bonferroni correction) to 1) compare vitamin B-12 status among reproductive groups at study-end, and 2) assess vitamin B-12 status changes from baseline to study-end within each reproductive group. Correlations between breast milk vitamin B-12 and circulating biomarkers of vitamin B-12 were also examined among the lactating women with use of Pearson’s correlation analysis (with log-transformed variables). Finally, to assess vitamin B-12 deficiency among study participants, the following cutoff values were used (33–36): serum vitamin B-12 <148 pmol/L; serum holotranscobalamin <35 pmol/L; serum MMA >271 nmol/L; and serum homocysteine >12 mmol/L. All statistical tests were two-sided, and significance was defined as \( P < 0.05 \). Analyses were performed with IBM SPSS
Statistics (version 20).
RESULTS

Characteristics of study participants

Baseline demographic and biochemical characteristics of pregnant, lactating, and control participants are shown in Table 3.1. No statistically significant differences were detected in age, ethnicity/race, BMI, and the distribution of TCN2 C776G genotype among reproductive groups. However, the percentage of multivitamin supplement use before study enrollment differed ($P = 0.001$) among reproductive groups with a higher percentage in pregnant and lactating women compared with control women.

At baseline, serum vitamin B-12 differed among reproductive groups ($P < 0.001$) with higher serum vitamin B-12 in lactating women than in control (+40%; $P < 0.001$) and pregnant (+45%; $P < 0.001$) women (Table 3.1). Serum holotranscobalamin concentrations also differed among reproductive groups ($P = 0.001$) with higher holotranscobalamin in lactating women than in control women (+57%; $P = 0.001$). In addition, serum MMA and homocysteine differed among reproductive groups ($P < 0.001$) with lower MMA and homocysteine concentrations in pregnant women than in control women (-25% and -44%, respectively; $P \leq 0.02$) and lactating women (-32% and -40%, respectively; $P < 0.001$). Lastly, urinary homocysteine differed among reproductive groups ($P < 0.001$) with higher excretion in pregnant women than in control (+33%; $P = 0.04$) and lactating (+94%; $P < 0.001$) women. Compared with control women, lactating women had lower urinary homocysteine (-31%; $P = 0.006$).
Table 3.1. Baseline characteristics of pregnant, lactating and control participants

<table>
<thead>
<tr>
<th></th>
<th>Pregnant (n=26)</th>
<th>Lactating (n=28)</th>
<th>Control (n=21)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>28 ± 3</td>
<td>29 ± 5</td>
<td>29 ± 5</td>
<td>0.82</td>
</tr>
<tr>
<td>Ethnicity/race, %</td>
<td></td>
<td></td>
<td></td>
<td>0.74</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>61.5</td>
<td>71</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Other2</td>
<td>38.5</td>
<td>29</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Baseline or prepregnancy BMI, kg/m²</td>
<td>23 [21-26]</td>
<td>25 [21-31]</td>
<td>24 [21-25]</td>
<td>0.46</td>
</tr>
<tr>
<td>TCN2 C776G genotype (rs1801198), %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC/CG/GG</td>
<td>57/38/5</td>
<td>35/46/19</td>
<td>26/52/22</td>
<td>0.20</td>
</tr>
<tr>
<td>Multivitamin supplement use before study entry, % yes</td>
<td>85³</td>
<td>75³</td>
<td>33³</td>
<td>0.001</td>
</tr>
<tr>
<td>Serum vitamin B-12, pmol/L</td>
<td>320 (288-356)⁶</td>
<td>463 (419-511)⁴</td>
<td>330 (290-376)⁵</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Serum holotranscobalamin, pmol/L</td>
<td>76 (66-89)⁶</td>
<td>96 (82-111)⁴</td>
<td>61 (50-73)⁵</td>
<td>0.001</td>
</tr>
<tr>
<td>Holotranscobalamin-to-vitamin B-12 ratio</td>
<td>0.24 (0.21-0.27)</td>
<td>0.21 (0.18-0.24)</td>
<td>0.19 (0.16-0.22)</td>
<td>0.10</td>
</tr>
<tr>
<td>Serum MMA, nmol/L</td>
<td>134 (116-154)⁶</td>
<td>198 (173-227)⁴</td>
<td>179 (153-209)⁵</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Serum homocysteine, µmol/L</td>
<td>3.7 (3.4-4.0)⁶</td>
<td>6.2 (5.8-6.7)⁴</td>
<td>6.6 (6.0-7.1)⁵</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Urinary homocysteine, µg/mg creatinine</td>
<td>0.60 (0.52-0.70)⁶</td>
<td>0.31 (0.27-0.36)⁴</td>
<td>0.45 (0.38-0.53)³</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

¹Values are means ± SDs, medians [IQRs], or geometric means (95% CIs), unless otherwise indicated. Differences between reproductive groups were analyzed by one-way ANOVA for normally distributed continuous variables (i.e., age and natural log-transformed vitamin B-12 biomarkers); Kruskal-Wallis tests for non-normally distributed continuous variable (i.e., BMI); and chi-square tests for categorical variables (i.e., ethnicity, multivitamin supplement use and TCN2 C776G genotype). Labeled values in a row without a common letter differ, P<0.05. Because of limited sample volume, serum vitamin B-12 was only measured in 24 pregnant, 27 lactating and 16 control women, and serum holotranscobalamin in all pregnant, 27 lactating and 18 control women. MMA, methylmalonic acid; TCN2, transcobalamin II.

²African American, Hispanic, Asian or other.
Vitamin B-12 status by reproductive state and its response through time

Serum vitamin B-12 concentration. Throughout the study, all participants had serum vitamin B-12 concentrations >148 pmol/L, indicating vitamin B-12 sufficiency of our study cohort. After covariate adjustment, reproductive state interacted with time ($P = 0.02$) to influence serum vitamin B-12 concentrations. Specifically, the concentration of serum vitamin B-12 did not change from baseline to study-end among pregnant and lactating women ($P = 1.0$), whereas it increased ($P = 0.009$) by ~23% among control women (Figure 3.1A). At study-end, lactating women had 26% higher serum vitamin B-12 than control women ($P = 0.04$) and 60% higher serum vitamin B-12 than pregnant women ($P < 0.001$). Although baseline serum vitamin B-12 did not differ between pregnant and control women, its concentration at study-end was 21% lower in pregnant (vs. control) women ($P = 0.02$).

Serum holotranscobalamin concentration. At baseline, serum holotranscobalamin concentrations of 1 pregnant woman and 2 control women were <35 pmol/L (range: 31.5–33.6 pmol/L). However, at study-end, serum holotranscobalamin concentrations among all participants were greater than this cutoff value. In covariate-adjusted analysis, reproductive state interacted with time ($P = 0.004$) to influence serum holotranscobalamin concentrations. Specifically, serum holotranscobalamin increased from baseline to study-end by ~19% among pregnant women ($P < 0.001$), 16% among lactating women ($P = 0.009$), and 42% among control women ($P < 0.001$) (Figure 3.1B). Although the baseline serum holotranscobalamin was greater in lactating than in control women ($P = 0.04$), its concentration at study-end did not differ by reproductive state ($P \geq 0.44$).

Ratio of serum holotranscobalamin to vitamin B-12. Reproductive state and time did not interact ($P = 0.51$) to influence the holotranscobalamin–to–vitamin B-12 ratio. Although the baseline
holotranscobalamin–vitamin B-12 ratio did not differ by reproductive state, its ratio at study-end was significantly higher in pregnant women than in control women (+33%; \( P = 0.04 \)) (Figure 3.1C). No differences were observed between lactating and control women nor between lactating and pregnant women (\( P \geq 0.26 \)).

**Serum MMA concentration.** At baseline, serum MMA concentrations of 1 pregnant, 6 lactating, and 2 control women were >271 nmol/L (range: 277–451 nmol/L), whereas at study-end, those of 1 pregnant and 2 control women exceeded this cutoff value (range: 281–347 nmol/L). After covariate adjustment, we observed a significant interaction (\( P < 0.001 \)) between reproductive state and time to influence serum MMA concentrations. Specifically, serum MMA increased from baseline to study-end by ~50% among pregnant women (\( P < 0.001 \)), but no changes were observed among control and lactating women (\( P \geq 0.13 \)) (Figure 3.1D). Although the baseline serum MMA was lower in pregnant women than in control and lactating women (\( P < 0.001 \)), its concentration at study-end did not differ by reproductive state (\( P = 1.0 \)).

**Serum homocysteine concentration.** Throughout the study, serum homocysteine concentrations among all participants were <12 µmol/L, indicating vitamin B-12 sufficiency of our study cohort. Similar to serum MMA, there was a significant interaction (\( P < 0.001 \)) between reproductive state and time in relation to serum homocysteine concentrations. Specifically, serum homocysteine increased from baseline to study-end by ~15% among pregnant women (\( P = 0.009 \)), but decreased by 16% among control women (\( P < 0.001 \)) and 17% among lactating women (\( P < 0.001 \)) (Figure 3.1E). Despite these changes, pregnant women had lower serum homocysteine than control (-22%; \( P < 0.001 \)) and lactating (-17%; \( P = 0.02 \)) women at study-end. No differences were observed between lactating and control women (\( P = 1.0 \)).

**Urinary homocysteine concentration.** After covariate adjustment, reproductive state tended to
interact \((P = 0.08)\) with time to influence urinary homocysteine excretion. Specifically, urinary homocysteine increased from baseline to study-end by \(~29\%\) among pregnant women \((P = 0.045)\), whereas no changes were observed among lactating and control women \((P = 1.0)\) (Figure 3.1F). At study-end, pregnant women excreted 75\% more homocysteine than control women \((P < 0.001)\) and 124\% more homocysteine than lactating women \((P < 0.001)\). Although the baseline urinary homocysteine excretion was lower in lactating women than in control women \((P = 0.009)\), its excretion at study-end did not differ between the 2 groups \((P = 0.39)\).
Serum vitamin B\textsubscript{12} (pmol/L)

Week

Control

Pregnant

Lactating

Serum holoTC, pmol/L

Week

Control

Pregnant

Lactating

Serum holoTC:vitamin B\textsubscript{12} ratio

Week

Control

Pregnant

Lactating

Serum MMA, nmol/L

Week

Control

Pregnant

Lactating

Serum homocysteine, µmol/L

Week

Control

Pregnant

Lactating

Urinary homocysteine, µg/mg creatinine

Week

Control

Pregnant

Lactating
Figure 3.1. Biomarkers of vitamin B-12 status in pregnant, lactating, and control women with equivalent vitamin B-12 intakes under controlled feeding conditions. All analyses were performed with use of linear mixed models; values are geometric means and 95% CIs. (A) Serum vitamin B-12; (B) serum holotranscobalamin; (C) the holotranscobalamin–to–vitamin B-12 ratio; (D) serum MMA; (E) serum homocysteine; and (F) urinary homocysteine. The study-end reflects week 12 for pregnant (n = 26) and control (n = 21) women and week 10 for lactating women (n = 28); because of limited sample volume, serum vitamin B-12 was only measured in 24 pregnant, 27 lactating, and 16 control women, and serum holotranscobalamin in all pregnant, 27 lactating, and 18 control women. Labeled endpoint means without a common letter differ, \( P < 0.05 \). *Different from baseline, \( P < 0.05 \) within a designated reproductive group. holoTC, holotranscobalamin; MMA, methylmalonic acid.
Correlations between breast milk vitamin B-12 and maternal biomarkers of vitamin B-12 status

The vitamin B-12 concentration of breast milk did not change ($P = 0.46$) from baseline [geometric mean (95% CI): 318 (227, 447) pmol/L] to study-end [298 (213, 419) pmol/L]. At baseline, there were significant, but modest, positive correlations of breast milk vitamin B-12 with maternal serum vitamin B-12 ($r = 0.48$, $P = 0.01$; Figure 3.2A) and holotranscobalamin ($r = 0.42$, $P = 0.03$; Figure 3.2B). Breast milk vitamin B-12 also tended to inversely correlate with maternal serum homocysteine ($r = 0.34$, $P = 0.08$; Figure 3.2C). Other than these findings, no statistically significant correlations were observed between maternal vitamin B-12 indicators (i.e., holotranscobalamin–to–vitamin B-12 ratio, serum MMA, and serum/urinary homocysteine) and breast milk vitamin B-12. At study-end, the positive correlation of breast milk vitamin B-12 remained significant with serum holotranscobalamin ($r = 0.40$, $P = 0.04$; Figure 3.2D), but not with serum vitamin B-12 concentrations ($r = 0.18$, $P = 0.36$).
A

In (baseline serum vitamin B-12, pmol/L)

In (baseline milk vitamin B-12, pmol/L)

\( r = 0.48 \)
\( P = 0.01 \)

B

In (baseline serum holoTC, pmol/L)

In (baseline milk vitamin B-12, pmol/L)

\( r = 0.42 \)
\( P = 0.03 \)

C

In (baseline serum Hcy, µmol/L)

In (baseline milk vitamin B-12, pmol/L)

\( r = -0.34 \)
\( P = 0.08 \)

D

In (study-end serum holoTC, pmol/L)

In (study-end milk vitamin B-12, pmol/L)

\( r = 0.40 \)
\( P = 0.04 \)
Figure 3.2. Correlations between women’s breast milk vitamin B-12 and serum vitamin B-12 (A), holotranscobalamin (B), and homocysteine (C) at baseline and between breast milk vitamin B-12 and serum holotranscobalamin at endpoint (D). All data were natural log (ln) transformed and analyzed by Pearson’s correlation analysis (n = 28); because of limited sample volume, serum vitamin B-12 and holotranscobalamin were only measured in 27 lactating women. Hcy, homocysteine; holoTC, holotranscobalamin.
DISCUSSION

To the best of our knowledge, this is the first controlled feeding study to assess the effects of reproductive status (i.e., pregnancy and lactation) on biomarkers of vitamin B-12. The following 2 main findings emerged: 1) reproductive state was associated with altered biomarkers of vitamin B-12 status, and 2) pregnant and lactating women may benefit from vitamin B-12 intakes exceeding current recommendations.

Third trimester pregnant women had significantly lower (~21%) serum vitamin B-12 than control women at study-end despite equivalent vitamin B-12 intakes. This reduction, as suggested by other studies (10, 13, 37, 38), may be attributable to normal physiologic consequences of pregnancy including hemodilution, hormonal changes, and/or vitamin B-12 transfer from mother to fetus. For example, throughout pregnancy, maternal plasma volume expands an average of 45% to meet the increased circulatory needs of maternal and fetal organs (39). This in turn can dilute serum vitamin B-12 concentrations among pregnant women and contribute to their lower vitamin B-12 concentrations. Notably, the concentration of serum holotranscobalamin did not differ between pregnant and control women, suggesting that greater amounts of vitamin B-12 are partitioned toward the biologically active form (i.e., holotranscobalamin) in this reproductive state. Indeed, compared with control women, pregnant women had a significantly higher holotranscobalamin–to–vitamin B-12 ratio at study-end, with a greater proportion of serum vitamin B-12 (i.e., ~30% vs. 20%) bound to transcobalamin. This metabolic alteration would be expected to augment the supply of vitamin B-12 to the fetus (40). Alternatively, but less likely, higher maternal holotranscobalamin concentrations may suggest impaired or limited placental uptake of holotranscobalamin.

Lactation was also found to alter vitamin B-12 status with significantly higher total serum
vitamin B-12 observed among lactating (vs. control) women throughout the study period. Given that serum holotranscobalamin did not differ between the 2 groups at study-end and that the vitamin B-12 bound to haptocorrin can be approximated by subtracting holotranscobalamin from vitamin B-12 concentrations (41), these data imply that lactating (vs. control) women maintain higher circulating vitamin B-12 in the form bound to haptocorrin. Although the physiologic role of vitamin B-12–haptocorrin complex is not fully understood, it may serve as a circulating storage form of vitamin B-12, which would be expected to attenuate loss of free vitamin B-12 (42, 43). However, given that haptocorrin is not taken up by the mammary epithelium (44), redistribution of vitamin B-12 from haptocorrin to transcobalamin may occur in the circulation and/or liver (a major storage site for vitamin B-12) when holotranscobalamin supply is insufficient. Additional studies are required to confirm these findings by directly measuring concentrations of vitamin B-12 bound to haptocorrin and to provide mechanistic insights into this metabolic adaptation.

Over the course of the study, serum MMA concentrations increased significantly in pregnant women but not in control and lactating women. Despite this pregnancy-induced increase, serum MMA did not differ among reproductive groups at study-end. Moreover, the increased serum MMA during pregnancy did not appear to indicate functional vitamin B-12 depletion because the mean serum MMA concentration remained <271 nmol/L (34). Previous reports (45, 46) have also suggested that an increase in serum or urinary MMA may not reflect low vitamin B-12 status during normal pregnancy. As such, serum MMA may not be a reliable biomarker of vitamin B-12 status among pregnant women.

In accordance with previous results (47–49), the concentration of serum homocysteine was lower among pregnant (vs. control) women throughout the study period. The present study
also reports, for the first time to our knowledge, that pregnant women excreted ~75% more urinary homocysteine than control women and ~124% more homocysteine than lactating women under equivalent vitamin B-12 intakes. Given that none of the reproductive groups were vitamin B-12 deficient, lower serum homocysteine and higher urinary homocysteine concentrations among pregnant women may arise in part from pregnancy-induced physiologic changes including hemodilution and higher glomerular filtration rate. Indeed, the concentrations of other relevant metabolites (e.g., cysteine and methionine) were lower in serum and higher in urine during pregnancy (data not shown), further supporting this hypothesis.

The mean daily consumption of ~8.6 µg vitamin B-12 (which was supplied through a mixed diet containing ~6 µg/d and a prenatal multivitamin supplement labeled to contain 2.6 µg/d) provided a vitamin B-12 intake ~3 times the RDA for pregnant women (2.6 µg/d) and lactating women (2.8 µg/d) (1). This amount of vitamin B-12 intake achieved adequate vitamin B-12 status among all reproductive groups, as indicated by biomarker concentrations that were above (serum vitamin B-12 and holotranscobalamin) or below (serum MMA and homocysteine) cutoff values. However, the mean concentration of serum vitamin B-12 among pregnant women (i.e., ~311 pmol/L) fell within the low- to mid-range of normal (i.e., 148–664 pmol/L) (50), suggesting that vitamin B-12 intakes approximating the RDA of 2.6 µg/d may not ensure acceptable vitamin B-12 status throughout this reproductive state.

Consumption of the study vitamin B-12 dose also yielded a significant increase in serum holotranscobalamin concentrations among all reproductive groups, which would enhance tissue access to this bioactive form. However, breast milk vitamin B-12 concentration did not change among lactating women in response to the study vitamin B-12 dose, despite a positive correlation between maternal serum holotranscobalamin (but not total serum vitamin B-12) and breast milk
vitamin B-12 both at baseline and study-end. Additional studies with multiple time points and methodology that enables measurements of the various forms of vitamin B-12 in breast milk are required to more fully understand the relation between maternal vitamin B-12 indicators and milk vitamin B-12 concentration.

The present study has 2 main limitations. First, no comparable group with a lower vitamin B-12 intake (e.g., the vitamin B-12 RDAs) was included in this study. Thus, additional doseresponse studies are needed to further assess the adequacy of the vitamin B-12 RDA during pregnancy and lactation. Second, the small sample size of the reproductive groups in this feeding study could limit detection of subtle differences between groups and across time.

In conclusion, metabolic adaptations that increase vitamin B-12 supply to the child occur during pregnancy and lactation. Consumption of the study vitamin B-12 dose (~3 times the RDA for pregnant and lactating women) improved circulating concentrations of the bioactive form of this essential nutrient, suggesting that women in these reproductive states may benefit from vitamin B-12 intakes exceeding current recommendations.
REFERENCES


3. Bor MV, Lydeking-Olsen E, Møller J, Nexø E. A daily intake of approximately 6 microg vitamin B-12 appears to saturate all the vitamin B-12-related variables in Danish postmenopausal women. Am J Clin Nutr 2006;83:52–8.


Supplemental Table S3.1. Estimated vitamin B-12 content of foods in seven-day rotational menu consumed by pregnant, lactating and control women for 10-12 weeks

<table>
<thead>
<tr>
<th>Day</th>
<th>Estimated daily vitamin B-12 content, µg/d</th>
<th>Breakfast</th>
<th>Lunch</th>
<th>Dinner</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2 eggs, scrambled (50g each)</td>
<td>Pesto sandwich:</td>
<td>Beef &amp; Cheese tacos:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 slices whole wheat toast (28g each)</td>
<td>2 slices whole wheat bread (28g each)</td>
<td>3 Corn tortillas (33g each)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 peach fruit cup (133g)</td>
<td>Pesto (15g)</td>
<td>Ground beef (105g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juice</td>
<td>Swiss Cheese (42g)</td>
<td>Cheddar cheese (30g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Romaine lettuce (20g)</td>
<td>Iceberg lettuce (20g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Celery sticks (30g)</td>
<td>Melon (112g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Carrot sticks (30g)</td>
<td>Milk (284g)</td>
</tr>
<tr>
<td>Monday</td>
<td>7.8</td>
<td>Waffle (130g)</td>
<td>Tuna sandwich:</td>
<td>Spaghetti:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 egg, hard-boiled (50g)</td>
<td>2 slices whole wheat bread (28g each)</td>
<td>Cooked pasta (220g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juice</td>
<td>Tuna, canned (56g)</td>
<td>Tomato sauce (220g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cheddar Cheese (30g)</td>
<td>Mushrooms (30g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Iceberg lettuce (30g)</td>
<td>Parmesan cheese (10g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mayonnaise (15g)</td>
<td>Mozzarella cheese (25g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Grapes (100g)</td>
<td>Milk (284g)</td>
</tr>
<tr>
<td>Tuesday</td>
<td>9.0</td>
<td>Fitness crunch cereal (80g)</td>
<td>Pastrami sandwich:</td>
<td>Vegetarian pizza:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Milk (284g)</td>
<td>2 slices whole wheat bread (28g each)</td>
<td>Dough (200g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 box of raisins (42g)</td>
<td>Pastrami (24g)</td>
<td>Tomato sauce (112g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 banana (118g)</td>
<td>Swiss cheese (23g)</td>
<td>Red peppers, jarred (60g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juice</td>
<td>Romaine lettuce (30g)</td>
<td>Mushrooms (30g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cucumber (80g)</td>
<td>Spinach (20g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mozzarella cheese (50g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Apple sauce (28g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Milk (284g)</td>
</tr>
<tr>
<td>Day</td>
<td>Number</td>
<td>Meal Description</td>
<td>Additional Ingredients</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Thursday</td>
<td>5.6</td>
<td>2 pancakes (100g each)</td>
<td>Egg salad sandwich:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Blueberries, frozen (50g)</td>
<td>2 slices whole wheat bread (28g each)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juice</td>
<td>1 egg, hard-boiled (50g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mayonnaise (20g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Romaine lettuce (45g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Celery sticks (30g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Carrot sticks (30g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Beef &amp; broccoli stir-fry:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Beef (160g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Broccoli, frozen (110g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rice (200g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Onions, frozen (13g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Melon (112g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Milk (284g)</td>
<td></td>
</tr>
<tr>
<td>Friday</td>
<td>4.9</td>
<td>1 whole wheat bagel (95g)</td>
<td>Bean Burrito:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 orange fruit cup (133g)</td>
<td>Black beans, canned (40g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juice</td>
<td>Cheddar cheese (30g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rice (30g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 large whole wheat tortilla (59g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lasagna:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 whole wheat lasagna noodles, dry (23g each)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tomato sauce (150g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ground beef (60g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cottage cheese (40g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mozzarella cheese (10g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Parmesan cheese (10g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Summer squash (40g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Zucchini (40g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Milk (284g)</td>
<td></td>
</tr>
<tr>
<td>Saturday</td>
<td>4.3</td>
<td>2 raspberry muffins (60g each)</td>
<td>Turkey Sandwich:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 banana (118g)</td>
<td>2 slices whole wheat bread (28g each)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juice</td>
<td>Turkey (25g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Provolone cheese (23g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Iceberg lettuce (20g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cucumber (80g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chicken Quesadilla:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chicken (65g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 large whole wheat tortillas(59g each)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cheddar cheese (150g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Corn, frozen (100g)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Milk (284g)</td>
<td></td>
</tr>
</tbody>
</table>
**Supplemental Table S3.1 (Continued)**\(^{1-3}\)

<table>
<thead>
<tr>
<th>Day</th>
<th>Estimated daily vitamin B-12 content, (\mu g/d)</th>
<th>Breakfast</th>
<th>Lunch</th>
<th>Dinner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunday</td>
<td>5.9</td>
<td>Oat granola cereal (80g)</td>
<td>Vegetable soup (396 g)</td>
<td>Goulash:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Milk (284g)</td>
<td>Corn muffin (60g)</td>
<td>Cooked pasta (200g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 box of raisins (42g)</td>
<td>Grapes (100g)</td>
<td>Beef (120g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 peach fruit cup (133g)</td>
<td></td>
<td>Onion, frozen (30g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juice</td>
<td></td>
<td>Red peppers, jarred (30g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tomatoes, canned (50g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tomato puree (100g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Potatoes, canned (35g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pineapple, canned (70g)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Milk (284g)</td>
</tr>
</tbody>
</table>

\(^{1}\)The vitamin B-12 content of the foods and beverages were estimated based on the US Department of Agriculture National Nutrient Database for Standard Reference, Release 22.  
\(^{2}\)Snacks including 170 g yogurt (vanilla, raspberry, or peach flavored) and 156 g V8 juice (Campbell Soup Company) were provided daily and contained 0.85 \(\mu g\) vitamin B-12 (which was added to the final vitamin B-12 content for each day).  
\(^{3}\)Juice choices included apple, cranberry-grape, or cranberry juice which did not contain vitamin B-12.
CHAPTER 4

Alcohol dehydrogenase 5 is a source of formate for de novo purine biosynthesis in HepG2 Cells*

ABSTRACT

**Background:** Formate provides one-carbon units for de novo purine and thymidylate (dTMP) synthesis and is produced via both folate-dependent and folate-independent pathways. Folate-independent pathways are mediated by cytosolic alcohol dehydrogenase 5 (ADH5) and mitochondrial aldehyde dehydrogenase 2 (ALDH2), which generate formate by oxidizing formaldehyde. Formate is a potential biomarker of B-vitamin–dependent one-carbon metabolism.

**Objective:** This study investigated the contributions of ADH5 and ALDH2 to formate production and folate-dependent de novo purine and dTMP synthesis in HepG2 cells.

**Methods:** *ADH5* knockout and ALDH2 knockdown HepG2 cells were cultured in folate-deficient [0 nM (6S) 5-formyltetrahydrofolate] or folate-sufficient [25 nM (6S) 5-formyltetrahydrofolate] medium. Purine biosynthesis was quantified as the ratio of [14C]-formate to [3H]-hypoxanthine incorporated into genomic DNA, which indicates the contribution of the de novo purine synthesis pathway relative to salvage synthesis. dTMP synthesis was quantified as the ratio of [14C]-deoxyuridine to [3H]-thymidine incorporation into genomic DNA, which indicates the capacity of de novo dTMP synthesis relative to salvage synthesis.

**Results:** The [14C]-formate-to-[3H]-hypoxanthine ratio was greater in *ADH5* knockout than in wild-type HepG2 cells, under conditions of both folate deficiency (+30%; *P* < 0.001) and folate sufficiency (+22%; *P* = 0.02). These data indicate that ADH5 deficiency increases the use of exogenous formate for de novo purine biosynthesis. The [14C]-deoxyuridine-to-[3H]-thymidine ratio did not differ between *ADH5* knockout and wild-type cells, indicating that ADH5 deficiency does not affect de novo dTMP synthesis capacity relative to salvage synthesis. Under folate deficiency, ALDH2 knockdown cells exhibited a 37% lower ratio of [14C]-formate to [3H]-
hypoxanthine ($P < 0.001$) compared with wild-type HepG2 cells, indicating decreased use of exogenous formate, or increased endogenous formate synthesis, for de novo purine biosynthesis.

**Conclusion:** In HepG2 cells, ADH5 is a source of formate for de novo purine biosynthesis, especially during folate deficiency when folate-dependent formate production is limited. Formate is also shown to be limiting in the growth of HepG2 cells.
INTRODUCTION

Folate, an essential B-vitamin, serves as a cofactor in the form of tetrahydrofolate (THF) polyglutamates to carry and activate one-carbon units for the de novo synthesis of purines and thymidylate (dTMP) and for the remethylation of homocysteine to methionine (Figure 4.1) (1). Disruptions in folate-mediated one-carbon metabolism (FOCM), which can arise from insufficient intake of folate and other nutrients (e.g., choline and vitamins B-12 and B-6) and/or genetic variants, have been linked to a higher risk of cancer, neurodegenerative diseases, and developmental anomalies (1).

FOCM occurs in the cytosol, mitochondria, and nucleus, all of which are interrelated through the exchange of one-carbon units from formate, serine, and glycine (2). Mitochondrial one-carbon metabolism plays a role in the generation of formate from the catabolism of serine, glycine, dimethylglycine, and sarcosine (Figure 4.1). Mitochondria-derived formate then enters the cytosol or nucleus and is incorporated into 10-formylTHF, which provides the C2 and C8 carbon for de novo purine biosynthesis. Alternatively, cytosolic 10-formylTHF may be reduced to 5,10-methyleneTHF, which can be used for de novo dTMP synthesis or irreversibly reduced to 5-methylTHF for homocysteine remethylation to methionine (Figure 4.1). Therefore, formate is an intermediate metabolite essential for one-carbon metabolism, its utilization and mitochondrial production are linked to folate status, and hence formate concentrations in serum have the potential to serve as a biomarker of folate status (3). Folate is essential both for formate synthesis and to sequester formate within the cell, because previous studies have reported increased plasma and urinary formate and decreased rates of formate production in rats and mice fed a folate-deficient diet compared with those fed a folate-replete diet (4–6).
Figure 4.1. Cytosolic and mitochondrial folate-mediated one-carbon metabolism and formate generation by ADH5 and ALDH2. ADH5, alcohol dehydrogenase 5; ALDH2, aldehyde dehydrogenase 2; DHF, dihydrofolate; DMG, dimethylglycine; dTMP, thymidylate; dUMP, deoxyuridine monophosphate; GSH, reduced glutathione; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine; THF, tetrahydrofolate.
In addition to formate synthesis through the folate-dependent mitochondrial pathways, formate can be produced through folate-independent pathways, one of which involves the oxidation of formaldehyde to formate (3). This can occur via the reaction mediated by the cytosolic glutathione and NAD+-dependent enzyme alcohol dehydrogenase 5 [ADH5, class III (alternative abbreviation, ADH3); also known as formaldehyde dehydrogenase]. Specifically, ADH5 oxidizes S-hydroxymethylglutathione, a molecule formed spontaneously from formaldehyde and glutathione, to S-formylglutathionine, which is further converted to formate (Figure 4.1) (7). Formaldehyde can also be oxidized to formate by a mitochondrial NAD+-dependent aldehyde dehydrogenase class II (ALDH2; Figure 4.1) (8). ALDH2 and ADH5 are ubiquitously expressed in various tissues, including liver, kidney, and brain, and are most abundant in the liver (9–12).

Although there has been increasing interest in the role of mitochondria-derived formate in the functioning of FOCM and as a biomarker of nutrient status, very little is known about the contribution of folate-independent sources of formate to one-carbon metabolism. The objective of the current study was to investigate the effects of ADH5 and ALDH2 activity on de novo purine and dTMP biosynthesis by inhibiting ADH5 or ALDH2 expression in human hepatocarcinoma (HepG2) cells.
METHODS

*Cell culture.* HepG2 cells were maintained in DMEM (Corning) with 10% (vol:vol) FBS (HyClone), 1% penicillin/streptomycin (Corning), and 4 mM L-glutamine (Corning) at 37°C and 5% CO₂. For all experiments, modified DMEM lacking glycine, serine, methionine, folate, choline, and all nucleosides/nucleotides was used with 10% dialyzed and charcoal-treated FBS, 1% penicillin/streptomycin, and 4 mM L-glutamine.

*Generation of ADH5 knockout HepG2 cells by CRISPR/Cas9.* The CRISPR single guide RNA (5’-TGAACATGGCGAACGAGGTA-3’) targeting exon 1 of human ADH5 (NM_000671) was cloned into the pSpCas9(BB)-2A-Puro CRISPR/Cas9 vector as previously described (13). Cells were transfected for 48 h by using the FuGene 6 transfection reagent (Promega) following the manufacturer’s instructions. The transfected cells were selected in the presence of 2 µg puromycin/mL (RPI). The efficiency of ADH5 knockout was verified by immunoblotting.

*Gene knockdown by small interfering RNA transfection.* Cells were transfected with either negative control small interfering RNA (siRNA; Qiagen) or FlexiTube GeneSolution (GS217) siRNA for ALDH2 (Qiagen) by using Lipofectamine RNAiMAX (Life Technologies) according to the manufacturer’s instructions. Cells were harvested 72 h after transfection. The efficiency of ALDH2 knockdown was verified by immunoblotting.

*Cellular total folate measurement.* Total folate concentrations in cells were quantified by using a *Lactobacillus casei* microbiological assay as previously described (14).

*Immunoblotting.* Cellular proteins were extracted and quantified as previously described (15). Proteins were resolved on 4–15% (vol:vol) gradient SDS-PAGE gels (Bio-Rad) and transferred to Immobilon-P PVDF membrane (Millipore). The membrane was blocked for 1 h at room temperature in 5% BSA in PBS with 0.2% Tween. Primary antibodies were diluted in 5% BSA
in PBS with 0.2% Tween and incubated overnight at 4°C. Secondary antibodies were diluted in 5% nonfat dry milk in PBS with 1% Nonidet P-40 (US Biologicals) and added to the membrane for 1 h at room temperature. ADH5 and ALDH2 were detected with a 1:1000 rabbit anti-ADH5 antibody and a 1:2000 rabbit anti-ALDH2 antibody, respectively (Proteintech Group), followed by a 1:5000 dilution of HRP-conjugated donkey anti-rabbit secondary antibody (Pierce). As loading controls, 1:1000 mouse anti-α-Tubulin antibody (Active Motif) and a 1:3000 mouse anti-α-Calpain antibody (Affinity BioReagents) were used followed by a 1:5000 dilution of HRP-conjugated goat anti-mouse secondary antibody (Pierce). The membrane was visualized by autoradiography after the addition of SuperSignal West Pico Chemiluminescent Substrate (Pierce).

**Purine biosynthesis assay.** Cells were seeded on 100-mm plates in modified DMEM lacking glycine, serine, and all nucleosides/nucleotides but supplemented with 200 μM methionine and 30 μM choline, with 25 nM (6S) 5-formylTHF (folate sufficiency) or without (6S) 5-formylTHF (folate deficiency). After 2 doublings, cells were plated in triplicate on 6-well plates and allowed to grow for another doubling in the same media but supplemented with 10 μM [14C]-formate and 1 nM [3H]-hypoxanthine (Moravek Biochemicals). Cells were harvested, and genomic DNA was isolated by using a High Pure PCR template preparation kit (Roche) with RNase A treatment according to the manufacturer’s instructions. Isotope incorporation into genomic DNA was quantified by using a Beckman LS6500 scintillation counter in dual disintegrations/minute mode (16). Data are shown as the ratio of [14C]-formate to [3H]-hypoxanthine, which indicates the incorporation of formate into DNA via the folate-dependent de novo purine synthesis pathway relative to the incorporation of hypoxanthine into DNA via the folate-independent purine salvage pathway.
**dTMP biosynthesis assay.** Cells were plated and grown in modified DMEM lacking glycine, serine, and all nucleosides/nucleotides but supplemented with 200 µM methionine and 30 µM choline, with 25 nM (6S) 5-formylTHF (folate sufficiency) or without (6S) 5-formylTHF (folate deficiency). After 2 doublings, cells were plated in triplicate on 6-well plates and allowed to grow for another doubling in the same media but supplemented with 2 µM $[^{14}\text{C}]$-deoxyuridine and 25 nM $[^{3}\text{H}]$-thymidine (American Radiolabeled Chemicals). $[^{14}\text{C}]$-Deoxyuridine is incorporated into DNA via the folate-dependent de novo pathway, whereas $[^{3}\text{H}]$-thymidine is incorporated into DNA via the salvage pathway. Total genomic DNA was isolated from the harvested cells, and the isotope incorporation was quantified as described above. Data are shown as the ratio of $[^{14}\text{C}]$-deoxyuridine to $[^{3}\text{H}]$-thymidine (17).

**Cell growth assay.** Cell proliferation was determined by using a colorimetric MTT [3-(4,5)-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] assay (18). Cells were plated in 96-well plates and grown in modified DMEM lacking glycine, serine, and all nucleosides/nucleotides but supplemented with 200 µM methionine, 30 µM choline, and 25 nM (6S) 5-formylTHF. The effect of formate supplementation on cell proliferation was determined by using the same media further supplemented with either 30 µM or 90 µM sodium formate. From 24 to ≤72 h, cell growth was measured by adding 20 µL of 2.5 g MTT reagent/L to each well followed by 4 h of incubation at 37°C in 5% CO2. The insoluble formazan product was resuspended in 100 µL DMSO, and $A_{570}$ was measured on a microplate reader (Epoch; BioTek).

**Statistical analysis.** Histograms and scatterplots of the residuals were used to assess normality and variance homogeneity. The effects of gene expression (wild-type compared with $ADH5$ knockout or wild-type compared with $ALDH2$ knockdown) were assessed by using $t$ tests. To examine the effects of gene expression and folate status as well as their interaction, a 2-factor
ANOVA was used with post hoc Bonferroni corrections. Linear mixed models with Bonferroni corrections were used to assess the effect of gene expression on cell growth over time. Data are shown as means ± SDs of 3–5 biological replicates per condition. All statistical tests were performed with IBM SPSS Statistics (version 20), and significance was defined as $P < 0.05$. 
RESULTS

ADH5 deficiency increases use of exogenous formate for de novo purine biosynthesis

HepG2 cells lacking ADH5 were generated by using CRISPR/Cas9 genome editing (13). Ablation of ADH5 protein expression was confirmed by immunoblotting (Figure 4.2A). The wild-type and ADH5 knockout HepG2 cells were cultured in folate-deficient or folate-sufficient culture medium containing $[^{14}{\text{C}}]$-formate and $[^3{\text{H}}]$-hypoxanthine. The ratio of $[^{14}{\text{C}}]$-formate to $[^3{\text{H}}]$-hypoxanthine in DNA serves as a measure of de novo purine synthesis efficiency relative to salvage purine synthesis.

Culturing cells in folate-deficient medium increased the ratio of $[^{14}{\text{C}}]$-formate to $[^3{\text{H}}]$-hypoxanthine in both ADH5 knockout and wild-type HepG2 cells ($P < 0.001$; Figure 4.2B). This effect was driven by increased incorporation of $[^{14}{\text{C}}]$-formate into DNA (Supplemental Figure S4.1). Given that folate is required for mitochondria-derived formate production, we hypothesized that loss of ADH5 protein expression may lead to greater incorporation of exogenous formate into DNA under folate-deficient culture conditions than under folate-sufficient culture conditions. Notably, there was a significant interaction ($P < 0.001$) between the gene expression (comparing wild-type with ADH5 knockout) and folate in the culture medium (folate deficiency compared with folate sufficiency) on the ratio of $[^{14}{\text{C}}]$-formate to $[^3{\text{H}}]$-hypoxanthine (Figure 4.2B). Specifically, the ratio of $[^{14}{\text{C}}]$-formate to $[^3{\text{H}}]$-hypoxanthine was significantly greater in ADH5 knockout than in wild-type HepG2 cells in both culture conditions, but with a greater increase under conditions of folate deficiency (+30%; $P < 0.001$) than with folate sufficiency (+22%; $P = 0.02$). In the folate-deficient condition, $[^{14}{\text{C}}]$-formate incorporation normalized to DNA content was 30% greater in ADH5 knockout compared with wild-type HepG2 cells ($P < 0.001$), whereas the incorporation of $[^3{\text{H}}]$-hypoxanthine into DNA
did not differ \((P = 0.89)\) between the cell lines (Supplemental Figure S4.1). In folate sufficiency, the incorporation of \(^{14}\text{C}\)-formate into DNA did not differ between \(ADH5\) knockout and wild-type HepG2 cells \((P = 0.52)\), whereas \(^{3}\text{H}\)-hypoxanthine incorporation into DNA decreased by 12\% in \(ADH5\) knockout (compared with wild-type) HepG2 cells \((P = 0.008)\). Overall, these data suggest that \(ADH5\)-mediated formate production contributes to de novo purine biosynthesis, especially during folate deficiency when mitochondrial formate production is limited.

The effect of both \(ADH5\) expression and exogenous folate availability on intracellular folate concentrations was determined. Intracellular folate concentrations were \(~94\%\) lower in folate-deficient compared with folate-sufficient medium in both \(ADH5\) knockout and wild-type HepG2 cells \((P \leq 0.004; \text{Figure 4.2C})\). However, no effect of \(ADH5\) gene expression was observed on intracellular folate concentrations \((P = 0.43; \text{\(P\)-interaction between gene expression and folate in culture medium = 0.54})\).
Figure 4.2. Purine biosynthesis in WT and ADH5 knockout HepG2 cells. (A) Silencing of ADH5 was confirmed by immunoblotting. (B) The ratio of $[^{14}\text{C}]$-formate to $[^{3}\text{H}]$-hypoxanthine indicates the incorporation of formate into DNA via the de novo purine synthesis pathway relative to the incorporation of hypoxanthine into DNA via the purine salvage pathway. Labeled means without a common letter differ, $P < 0.05$. (C) Total cellular folate concentrations. *Different from folate-sufficient cells, $P < 0.05$. Data were analyzed by using 2-factor ANOVA with Bonferroni corrections. Values are means ± SDs of 3 biological replicates per condition. ADH5, alcohol dehydrogenase 5; DPM, decays per minute; KO, knockout; WT, wild-type.
ADH5 deficiency does not affect de novo dTMP synthesis

The effects of ADH5 silencing on dTMP synthesis in folate-deficient and folate-sufficient conditions were investigated. The ratio of [\(^{14}\)C]-deoxyuridine (an indicator of de novo Dtmp synthesis) to [\(^{3}\)H]-thymidine (an indicator of salvage dTMP synthesis) in DNA did not differ between the wild-type and ADH5 knockout HepG2 cells \((P = 0.14)\) independent of folate conditions \((P\text{-interaction} = 0.85)\) (Figure 4.3). When grown under folate-deficient compared with folate-sufficient conditions, both wild-type and ADH5 knockout HepG2 cells had a significantly lower ratio of [\(^{14}\)C]-deoxyuridine to [\(^{3}\)H]-thymidine \((P < 0.001)\), indicating that folate deficiency may upregulate the salvage pathway to meet cellular dTMP requirements.

**Figure 4.3.** dTMP biosynthesis in WT and ADH5 knockout HepG2 cells. The ratio of [\(^{14}\)C]-deoxyuridine to [\(^{3}\)H]-thymidine indicates the relative contribution of the de novo pathway to the salvage pathway for dTMP synthesis. Data were analyzed by using 2-factor ANOVA with Bonferroni corrections. Values are means ± SDs of 3 biological replicates per condition. *Different from folate-sufficient cells, \(P < 0.05\). ADH5, alcohol dehydrogenase 5; DPM, decays per minute; KO, knockout; WT, wild-type.
Reduced expression of ALDH2 decreases use of exogenous formate for de novo purine biosynthesis

To determine the effect of ALDH2 on purine biosynthesis, ALDH2 expression was reduced by using ALDH2-targeting siRNA transfection and verified by immunoblotting (Figure 4.4A). The wild-type and ALDH2 knockdown HepG2 cells were cultured in folate-deficient medium with $[^{14}\text{C}]$-formate and $[^{3}\text{H}]$-hypoxanthine. The ratio of $[^{14}\text{C}]$-formate to $[^{3}\text{H}]$-hypoxanthine was 37% lower in ALDH2 knockdown HepG2 cells compared with wild-type cells ($P < 0.001$; Figure 4.4B). ALDH2 knockdown HepG2 cells exhibited decreases in the incorporation of both $[^{14}\text{C}]$-formate (-82%; $P < 0.001$) and $[^{3}\text{H}]$-hypoxanthine (-72%; $P = 0.002$) into DNA (Supplemental Figure S4.2). Overall, these data suggest that reduced ALDH2 expression decreases the use of exogenous formate for de novo purine biosynthesis, indicating increased endogenous production of formate.
Figure 4.4. Purine biosynthesis in WT and ALDH2 knockdown HepG2 cells cultured in folate-deficient medium. (A) ALDH2 knockdown was confirmed by immunoblotting. (B) The ratio of \[^{14}\text{C}\]-formate to \[^{3}\text{H}\]-hypoxanthine indicates the incorporation of formate into DNA via the de novo purine synthesis pathway relative to the incorporation of hypoxanthine into DNA via the purine salvage pathway. Data were analyzed by using a \( t \) test. Values are means ± SDs of 3 biological replicates per condition.

*Different from WT, \( P < 0.05 \). ADH5, alcohol dehydrogenase 5; ALDH2, aldehyde dehydrogenase 2; DPM, decays per minute; KD, knockdown; KO, knockout; WT, wild-type.

Formate is limiting for HepG2 cell growth

The growth of wild-type, ALDH2 knockdown, and ADH5 knockout HepG2 cells in the presence of exogenous formate was investigated by supplementing culture medium with sodium formate. No significant interaction between gene expression (wild-type compared with \( \text{ADH5} \) knockout) and time was detected (\( P = 0.3 \)), indicating that the growth rate of \( \text{ADH5} \) knockout HepG2 cells did not differ from the wild-type cells (Supplemental Figure S4.3). However, ALDH2 knockdown HepG2 cells exhibited significantly lower growth rates relative to wild-type cells over 72 h [\( P \)-interaction between gene expression (wild-type compared with ALDH2 knockdown) and time < 0.001], indicating that reduced ALDH2 expression inhibits cell
proliferation. The ALDH2 knockdown HepG2 cells exhibited 23% and 31% lower cell viability at 48 and 72 h, respectively, compared with wild-type HepG2 cells ($P < 0.001$).

There was a significant interaction between gene expression (wild-type HepG2 cells compared with $ADH5$ knockout or ALDH2 knockdown HepG2 cells) and formate supplementation on cell proliferation ($P \leq 0.004$; Table 4.1). Specifically, the addition of formate (30 or 90 µM) in culture medium increased the growth of wild-type HepG2 cells at 48 and 72 h ($P < 0.001$); at 72 h, the wild-type cells supplemented with 90 µM formate exhibited a 30% increase in their growth compared with those without formate supplementation. However, the addition of formate did not affect the proliferation of ALDH2 knockdown or $ADH5$ knockout HepG2 cells ($P \geq 0.82$).
<table>
<thead>
<tr>
<th></th>
<th>Formate, µM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24h</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td><strong>WT</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.00 (± 0.02)</td>
</tr>
<tr>
<td><strong>ALDH2 knockdown</strong></td>
<td>0.85 (± 0.02)</td>
</tr>
<tr>
<td><strong>ADH5 knockout</strong></td>
<td>1.43 (± 0.04)</td>
</tr>
</tbody>
</table>

<sup>1</sup> Values are means ± SDs of 5 biological replicates per condition. Values were normalized to the A<sub>570</sub> value of WT HepG2 cells cultured in the medium without formate at 24 h. Data were analyzed by using a linear mixed model with Bonferroni corrections. Data from the formate 0-µM condition were presented in Supplemental Figure 4.3 to show the difference in the growth rates between the cell lines. Within each time point, labeled means in a row without a common letter differ, P < 0.05. ADH5, alcohol dehydrogenase 5; ALDH2, aldehyde dehydrogenase 2; WT, wild-type.
DISCUSSION

Formate provides one-carbon units for the de novo synthesis of purines and dTMP and for the remethylation of homocysteine to methionine. It also plays an important role in embryonic development, as reported by previous studies showing that formate has a protective effect on neural tube closure defects in a mouse model (3, 19, 20). Formate can be produced via both folate-dependent mitochondrial pathways and folate-independent pathways, but the relative contributions of these pathways to formate production and utilization are unknown (3). One of the folate-independent formate-generating pathways is mediated by NAD+-dependent cytosolic ADH5 and mitochondrial ALDH2, which function in the oxidation of formaldehyde to formate (Figure 4.1). The current study investigated whether ADH5 and ALDH2 enzymatic reactions contribute to the generation of endogenous formate for FOCM.

This study provides evidence that ADH5 is a meaningful source of formate for de novo purine biosynthesis. We found a higher ratio of $[{^{14}}\text{C}]$-formate to $[{^3}\text{H}]$-hypoxanthine in ADH5 knockout than in wild-type HepG2 cells, indicating that ADH5 knockout HepG2 cells incorporated higher amounts of exogenous formate into de novo purine biosynthesis than did wild-type cells. In addition, the use of exogenous formate for de novo purine synthesis in ADH5 knockout HepG2 cells increased during folate deficiency, presumably due to decreased mitochondria-derived formate generation. These findings are consistent with a previous study that showed that rats fed a folate-deficient diet exhibited a 44% reduction in the rate of endogenous formate production compared with those fed a folate-replete diet (5). Moreover, in liver mitochondria isolated from folate-deficient rats, formate production from choline metabolites (dimethylglycine and sarcosine) increased, which may be attributable to formaldehyde production, as suggested by the authors (5). Specifically, in folate deficiency,
dimethylglycine and sarcosine are sources of one-carbon units through the reactions mediated by dimethylglycine dehydrogenase and sarcosine dehydrogenase, thereby generating formaldehyde (21). However, HepG2 cells do not metabolize sarcosine or dimethylglycine (data not shown). Alternatively, endogenous formaldehyde can be generated as a byproduct of the enzymatic demethylation reactions, including histone, RNA, and DNA demethylation (22–24). These findings, as well as the results of our study, support a role for ADH5 in formate production through the oxidation of formaldehyde during folate deficiency. ADH5 also makes contributions to the formate pool in states of folate sufficiency. In this study, the ratio of $[^{14}\text{C}]$-formate to $[^{3}\text{H}]$-hypoxanthine was higher in the $ADH5$ knockout than in the wild-type HepG2 cells when they were cultured in the presence of 25 nM (6S) 5-formylTHF. Overall, our findings show a role for the folate-dependent and folate-independent production of formate for de novo purine biosynthesis.

Notably, the effect of ADH5 deficiency was different between de novo purine and dTMP synthesis. The ratio of $[^{14}\text{C}]$-deoxyuridine to $[^{3}\text{H}]$-thymidine in DNA did not differ between $ADH5$ knockout and wild-type HepG2 cells, indicating that dTMP synthesis is not compromised by ADH5 deficiency. De novo dTMP synthesis occurs at the sites of DNA synthesis in the nucleus (25), whereas de novo purine biosynthesis occurs in the cytoplasm, where it requires the formation of a multienzyme complex referred to as a purinosome (26). Although the underlying mechanisms need further elucidation, the differences in the effect of ADH5 deficiency between purine and dTMP synthesis suggest that the source of the nuclear formate pool may be different from that of the cytoplasmic formate pool. Given that ADH5 is localized in the cytosol, it may contribute to the formate pool in the cytosol, but not in the nucleus, for de novo purine biosynthesis.
ALDH2 knockdown HepG2 cells exhibited a 37% reduction in the ratio of $[^{14}C]$-formate to $[^{3}H]$-hypoxanthine compared with wild-type cells, indicating that reduced ALDH2 expression decreases the use of exogenous formate for de novo purine biosynthesis. These results indicate that ALDH2 knockdown HepG2 cells may upregulate endogenous formate production from the other sources as a compensatory response to ALDH2 deficiency. Alternatively, ALDH2 deficiency may enhance the conversion of formaldehyde to formate. However, given that the mitochondrial ratio of NAD$^+$ to NAD(H) ranges between 7 and 8 (27), the latter scenario that ALDH2 reduces formate seems unlikely. Overall, the differential effects on de novo purine biosynthesis between $ADH5$ knockout (increased exogenous formate use) and ALDH2 knockdown (decreased exogenous formate use) cells suggest that cytosolic ADH5 and mitochondrial ALDH2 may have distinct roles in providing formate for de novo purine biosynthesis.

ALDH2 knockdown HepG2 cells exhibited significantly slower growth rates relative to wild-type cells, indicating that reduced ALDH2 expression inhibits cell proliferation. Supplementation with formate in culture medium did not rescue the growth of these cells, suggesting that the observed growth inhibition is unlikely to be associated with formate production and/or utilization. As evidenced by a previous study (28), decreased proliferation of ALDH2 knockdown cells may be due to cell cycle arrest and enhanced apoptosis caused by elevated concentrations of reactive oxygen species and toxic aldehyde due to reduced ALDH2 activity. Interestingly, formate supplementation stimulated the growth of wild-type HepG2 cells, indicating that formate is limiting for the growth of these cells. This finding is consistent with a previous study that showed that methyleneTHF dehydrogenase 1 is essential for cell growth (29). Taken together, the results suggest that formate availability may be limiting for cell growth in
some cells. In conclusion, this study shows that endogenous formate produced by ADH5 is used in de novo purine biosynthesis and its contribution to the formate pool is enhanced in folate deficiency.
REFERENCES


9. Chen CH, Ferreira JC, Gross ER, Mochly-Rosen D. Targeting aldehyde dehydrogenase 2:
new therapeutic opportunities. Physiol Rev 2014;94:1–34.


127


Supplemental Figure S4.1. Purine biosynthesis in wild-type and ADH5 KO HepG2 cells. (A) The incorporation of \(^{14}\text{C}\)-formate into DNA were normalized to DNA content. (B) The incorporation of \(^{3}\text{H}\)-hypoxanthine into DNA were normalized to DNA content. Data were analyzed using a 2-way ANOVA with Bonferroni corrections. Values are shown as means ± SDs of three biological replicates per condition. Labeled means without a common letter differ, \(P < 0.05\). ADH5, alcohol dehydrogenase 5; DPM, decays per minute; KO, knockout; WT, wild-type.
Supplemental Figure S4.2. Purine biosynthesis in wild-type and ALDH2 KD HepG2 cells cultured in folate-deficient medium. The incorporation of $[^3]$H-hypoxanthine and $[^1]$C-formate into DNA was normalized to DNA content. Data were analyzed using $t$-test. Values are shown as means ± SDs of three biological replicates per condition. *Different from WT, $P < 0.05$. ALDH2, aldehyde dehydrogenase 2; DPM, decays per minute; KD, knockdown; WT, wild-type.
Supplemental Figure S4.3. Effect of ALDH2 KD and ADH5 KO on cell growth. Cell viability was determined in WT, ALDH2 KD, and ADH5 KO HepG2 cells. Data were analyzed using a linear mixed model with Bonferroni corrections. Values were normalized to the A_{570} of WT HepG2 cells at 24h and are shown as means ± SDs of five biological replicates per condition. Labeled means without a common letter differ, \( P < 0.05 \). ADH5, alcohol dehydrogenase 5; ALDH2, aldehyde dehydrogenase 2; KD, knockdown; KO, knockout; WT, wild-type.
AFTERWORD

This dissertation research aims to advance understanding of the impacts of folate, choline and vitamin B12 on the functioning of one-carbon metabolism and their relationship with colorectal cancer risk and reproductive state. This was achieved through a combination of human participant studies, laboratory-based molecular studies and a systematic review that integrate the areas of nutrition, epidemiology, molecular biology and public health. The major findings and potential implications are discussed below.

Relationship between folate status and DNA methylation is different before and after the U.S. folic acid fortification

This study investigated the association between folate status and DNA methylation in samples from postmenopausal women of the WHI-OS cohort collected before and after mandatory folic acid fortification. The results of this study indicate that the relationship between folate and DNA methylation differs across fortification periods with a positive relationship in the pre-fortification period but an inverse relationship in the post-fortification period. This suggests that the overall association between folate status and DNA methylation may follow an inverted U-shaped curve, and additional studies are warranted to clarify the potential health outcomes. If there are adverse health outcomes related to the inverse relationship observed in the era of folic acid fortification, folic acid supplement use may not be advisable among postmenopausal women in the U.S. or other countries with mandated folic acid fortification programs.

Alterations in choline metabolism are associated with higher risk of colorectal cancer

This was the first study to examine associations between plasma biomarkers of choline
metabolism and colorectal cancer risk among postmenopausal women in the U.S. The major findings indicate a positive association between plasma TMAO and rectal cancer risk, suggesting plasma TMAO as a potential biomarker for rectal cancer risk. Given that TMAO is a gut bacteria-derived metabolite, the positive association may be related to abnormal changes in colonic bacteria, which could occur in disease development. Interestingly, the association between TMAO and colorectal cancer risk was modified by vitamin B12 status with a positive association observed in women with lower plasma vitamin B12 but not in those with higher plasma vitamin B12. Overall, these findings represent evidence for correlations between nutrients, gut microbiome and colorectal cancer pathogenesis.

*Vitamin B12 status differs among pregnant, lactating and control women with equivalent nutrient intakes*

This study compared vitamin B12 status response among pregnant, lactating and control (nonpregnant, nonlactating) women who consumed equivalent vitamin B12 intakes. The findings of this study indicate that pregnancy is associated with altered biomarkers of vitamin B12 status with enhanced bioavailability of vitamin B12, which may augment the supply of vitamin B12 to the fetus. Specifically, a higher ratio of holotranscobalamin (bioactive form of vitamin B12) to total vitamin B12 was observed in pregnant (vs. control) women, suggesting that greater amounts of vitamin B12 are partitioned toward the biologically active form during pregnancy. This study also found that consumption of the study vitamin B12 dose (~3 times the RDA) yielded a significant increase in serum holotranscobalamin among all reproductive groups, warranting further investigation of clinical benefit, if any, of vitamin B12 intakes exceeding current recommendations.
**Alcohol dehydrogenase 5 is a source of formate for de novo purine biosynthesis in HepG2 cells**

The objective of this study was to investigate the effect of folate-independent formate-generating pathways mediated by alcohol dehydrogenase 5 (ADH5) and aldehyde dehydrogenase 2 (ALDH2) on one-carbon metabolism. This study demonstrates for the first time that ADH5 is a source of formate for *de novo* purine biosynthesis, especially during folate deficiency when folate-dependent formate production is limited. This study advances current understanding of a role for the folate-independent formate synthesis in the functioning of one-carbon metabolism, and further investigation is warranted to determine its impacts in different tissues and in whole animals.

**A systematic review: Provision of folic acid for reducing arsenic toxicity in arsenic-exposed children and adults**

This ongoing systematic review using the Cochrane methodology aims to summarize and evaluate the evidence on the effects of folic acid provision on arsenic-related health outcomes and biomarkers of arsenic and folate status in arsenic-exposed populations. Given that chronic arsenic exposure is a public health burden and that folate functions in detoxifying arsenic through arsenic methylation, this review will help determine whether folic acid interventions are efficacious in reducing arsenic toxicity and address their public health implications. The results of this review will also serve to highlight the direction and need for future research.
APPENDIX A

Provision of folic acid for reducing arsenic toxicity in arsenic-exposed children and adults:

a systematic review*

*Bae S, Kamynina E, Farinola AF, Caudill M, Stover PJ, Cassano PA, Berry R, Peña-Rosas JP.

Provision of folic acid for reducing arsenic toxicity in arsenic-exposed children and adults.

*Cochrane Database of Systematic Reviews 2017, Issue 5. Art. No.:CD012649. DOI:
10.1002/14651858.CD012649.
BACKGROUND

Description of the condition

Burden of chronic arsenic exposure

Arsenic is a common environmental toxin and exists in both organic and inorganic forms. In general, organic arsenic is considered to be less harmful than inorganic arsenic; however, depending on the chemical form, some organic compounds are toxic (ATSDR 2007) and also may undergo degradation giving rise to bioavailable inorganic arsenic species (Chávez-Capilla 2016). Acute poisoning by arsenic is rare, but low-level chronic exposure in humans through contaminated drinking water is common throughout the world. The World Health Organization (WHO) estimates that more than 140 million persons from more than 70 countries, including Bangladesh, China, India, Chile, Nepal and areas of the US, are chronically exposed to arsenic-contaminated water at levels exceeding the recommended safe concentration of 10 μg/L (Ahmed 2006; Chiou 2001; Ghosh 2013; Mazumder 2010; Naujokas 2013; Nielsen 2010; Rodríguez-Lado 2013; Sanders 2012; Smith 2000; Smith 2002; WHO 2008; Yu 2007). The primary source of contaminated drinking water is groundwater exposed to naturally occurring arsenic-rich geological formations (IARC 2012; Kim 2011). Rice, a staple food for half of the world’s population, can contain high levels of arsenic, which is derived from the soil of paddy fields (Gilbert-Diamond 2011; Ma 2008; Melkonian 2013; Sohn 2014; Stone 2008; Zavala 2008a; Zavala 2008b). Water contamination remains an issue despite mitigation efforts, and food grown or produced (or both) in arsenic-rich environments contributes to human arsenic exposure (Banerjee 2013; Carignan 2016; Davis 2012; Karagas 2016; Kippler 2016; Rose 2007). A high amount of arsenic, in both organic and inorganic forms, may also be present in some species of seaweed (Almela 2002; Brandon 2014; Khan 2015; Rose 2007), and in fish, shellfish and other
types of seafood (Choi 2010; Seo 2016). The average daily intake of total arsenic (including both organic and inorganic forms) from food and beverages in US and European populations has been estimated to be in the range of 20 to 300 μg/day (deCastro 2014; EFSA 2009; EFSA 2014; IARC 2012; IPCS 2001; Kurzius-Spencer 2014; Lynch 2014).

Health consequences associated with arsenic exposure in adults and children

Arsenic exposure affects almost every organ system in the body, including the brain. Chronic exposure to arsenic has been associated with neurotoxicity in adults as well as increased risk of cancer, diabetes, cardiovascular disease, kidney disease, anaemia and skin disease (ATSDR 2007; Axelson 1980; Cohen 2013; Ettinger 2009; Farzan 2013; IARC 1980; IARC 2012; Moon 2013; National Toxicology Program 2014; Naujokas 2013; Navas-Acien 2008; Prakash 2016; Sidhu 2015; Vahidnia 2007). Arsenic exposure through drinking water has also been linked to excess adult mortality (Argos 2010; Sohel 2009; Wu 1989; Yuan 2007; Yuan 2010) and adverse pregnancy outcomes, including preterm delivery, stillbirth, spontaneous abortion and low birth weight (Ahmad 2001; Hopenhayn-Rich 2000; Huyck 2007; Kile 2016; Laine 2015; Milton 2005; Rahman 2007). There is limited evidence from human studies suggesting a link between arsenic exposure and risk of neural tube defects (DeSesso 2001; Kwok 2006; Mazumdar 2015a; Wu 2011); furthermore, animal studies have demonstrated that a high dose of arsenic induces embryotoxicity, including neural tube defects and other congenital anomalies (Chaineau 1990; Han 2011; Hill 2008; Hill 2009; Morrissey 1983; Wlodarczyk 2001).

Exposure to arsenic in drinking water during early childhood or in utero is reported to increase subsequent mortality in young adults from lung cancer and other lung diseases (Smith 2006). Children have a higher metabolic rate than adults to support growth and development, which leads to a greater exposure to arsenic and a greater sensitivity to the adverse effects of
arsenic exposure (Bearer 1995). In the first six months of life, children drink seven times more water per kilogram of body weight than adults. Children aged between one and five years consume three to four times more food per kilogram of body weight than adults (Bearer 1995). Because of these metabolically-driven differences, exposure to arsenic in children younger than three years of age is estimated to be two to three-fold higher than adults (EFSA 2009; EFSA 2014; Hojsak 2015; Ljung 2011; Meharg 2008; Rintala 2014; Signes-Pastor 2016). Higher exposure to arsenic in drinking water and/or higher blood and urinary arsenic concentrations have been inversely associated with intellectual function and neuropsychological development in children (Dong 2009; Rodríguez-Barranco 2016; Tolins 2014; Tsuji 2015; von Ehrenstein 2007) and adolescents (Tsai 2003).

**Biomarkers of arsenic exposure**

Arsenic levels measured in blood and urine are reliable biomarkers of arsenic exposure and status (ATSDR 2007; Hall 2006). In general, blood arsenic and urinary arsenic levels greater than (> 1 μg/L and 100 μg/L respectively are considered abnormal (ATSDR 2007), while some studies suggest that health risks of arsenic exposure may be associated with total urinary levels > 50 μg/L (Tseng 2005; Valenzuela 2005; WHO 2001). Given that blood and urinary arsenic levels reflect short-term exposure (i.e. hours to days) to arsenic, these are considered as biomarkers amenable to modifications driven by interventions (e.g. dietary change) (Hall 2006). Arsenic levels measured in hair and nails reflect more prolonged exposure to arsenic because arsenic accumulates in these slow growing tissues (Hall 2006).

The toxicity of arsenic is influenced by its chemical form. Inorganic arsenic species found in ground water are mainly arsenites (\( \text{iAs}^{\text{III}} \)) and arsenates (\( \text{iAs}^{\text{V}} \)). Inorganic arsenic can be modified by the addition of one, two or three methyl (CH\(-\)) groups in a process known as
methylation, which leads to the organic forms known as monomethyl-, dimethyl-, or trimethyl-arsenicals, respectively. These forms are also referred to as monomethylarsonic acid (MMA\textsuperscript{V}), monomethylarsonous acid (MMA\textsuperscript{III}), dimethylarsonous acid (DMA\textsuperscript{III}), dimethylarsonic acid (DMA\textsuperscript{V}) and tri-methyl-arsine oxide. The most toxic arsenicals are iAs\textsuperscript{III}, MMA\textsuperscript{III} and DMA\textsuperscript{III}, followed by iAs\textsuperscript{V}, MMA\textsuperscript{V} and DMA\textsuperscript{V} (ATSDR 2007; Styblo 2000).

Population groups exposed to arsenic mainly via drinking water typically excrete 10% to 30% as inorganic arsenicals (iAs\textsuperscript{III} and iAs\textsuperscript{V}), 10% to 20% as monomethylated arsenicals (MMA\textsuperscript{III} and MMA\textsuperscript{V}), and 60% to 70% as dimethylated arsenicals (DMA\textsuperscript{III} and DMA\textsuperscript{V}) (Vahter 2000). Because methylated arsenicals are more easily excreted through urine, decreased methylation of arsenic is associated with its increased retention in the body, which leads to increased toxicity. For example, compared to women, men show higher urinary monomethyl arsenicals, which indicate incomplete methylation of arsenic, and men also show more frequent skin lesions (Lindberg 2008). In arsenic-exposed population studies, participants with skin lesions (versus those without skin lesions) had a higher proportion of monomethyl arsenicals in urine, which is consistent with a lower arsenic methylation capacity (Li 2011; Valenzuela 2005; Zhang 2014). Arsenic methylation patterns in children differ from those in adults (Concha 1998; Fängström 2009; Skröder Löveborn 2016). In preschool children, urinary total arsenic and monomethyl-As\textsuperscript{V} (MMA\textsuperscript{V}) percentage were shown to be positively associated with the risk of developmental delay (Hsieh 2014).

Several factors affect the extent of arsenic methylation in the body. Dietary factors that function in cellular methylation pathways can affect methylation of arsenic and therefore its excretion. Among these, folate, a B vitamin, has been suggested as an important dietary source for facilitating arsenic methylation and excretion (Carlin 2016). Thus, dietary interventions to
increase folate are a potential means to reduce arsenic toxicity and prevent arsenic associated diseases. Other physiological factors that may affect excretion of arsenic include gut microbiome (Pinyayev 2011; Rubin 2014), gender (Jansen 2016; Lindberg 2008), pregnancy (Gardner 2011; Gardner 2012), and body mass index (Gribble 2013).

**Description of the intervention**

Folate is a general term for the water-soluble vitamin B₉ naturally present in foods, which humans are not able to synthesize in vivo and which therefore has to be obtained from dietary sources. Folate serves as a carrier for methyl groups required for biochemical reactions within cells, including methylation of arsenic. Naturally occurring folates exist in many chemical forms and are unstable, while folic acid is the stable, synthetic, oxidized form used in supplements and fortified food (Fox 2008). When used as a dietary supplement or fortificant, folic acid is mostly metabolized to the metabolically active, natural forms of folate, including 5-methyltetrahydrofolate, which is the form found in blood (Pfeiffer 2015; Stover 2004). Blood folate concentration can be reported as red blood cell (RBC) folate or serum/plasma folate. Serum blood folate levels are the earliest indicators of recent exposure and reflect recent dietary intake (short-term status). RBC folate is a sensitive indicator of folate status in the preceding 120 days (Bailey 2015). Elevated serum/plasma homocysteine is a sensitive biomarker of folate deficiency; however, elevations in homocysteine are not specific given that they increase in other B-vitamin deficiencies, such as B12 deficiency, and can be affected by other factors such as renal insufficiency and drug treatments (Bailey 2015). Nonetheless, plasma homocysteine is highly responsive to intervention with folate, either alone or in combination with the other methyl donors involved in one-carbon metabolism such as betaine, choline, and vitamins B₂, B₆,
and B₁₂.

Folate is critical for supporting rapid fetal growth and development and thus is especially important for women who may become pregnant. Failure of the neural tube to close during early embryonic development results in serious congenital anomalies collectively referred to as neural tube defects (Beaudin 2009; Botto 1999; Greene 2014). A Cochrane systematic review has showed that folic acid supplementation in women planning to become pregnant decreased the incidence and recurrence of neural tube defects in the fetus (De-Regil 2015). Currently, the US Preventive Services Task Force recommends that all women planning, or capable of, pregnancy take a daily supplement containing 400 to 800 µg of folic acid (USPSTF 2009) with a goal to reduce the occurrence of neural tube defects. To help achieve these recommended folate intake levels, the U.S. Food and Drug Administration (FDA) mandated fortification of the food supply through enriched cereal grain products (e.g., bread, pasta, rice fortified with 140 µg folic acid per 100 g grain). Population-wide folic acid fortification has been introduced in 84 countries and has been effective in reducing the rates of live-born infants with spina bifida (Atta 2016; Cordero 2015; FFI 2016; Grosse 2016, Rader 2006).

Folic acid supplementation in women of child-bearing age, and population-wide fortification of staple foods are currently used as public health interventions to reduce the rate of neural tube defects. Given the function of folate to detoxify arsenic through arsenic methylation, this review evaluates and summarizes the evidence on whether similar folic acid interventions are efficacious in reducing the public health burden of arsenic-associated health outcomes in arsenic-exposed adults and children, including all ages and gender groups.

How the intervention might work
An important folate-dependent reaction is the conversion of homocysteine to methionine (Caspi 2016). In this reaction, the enzyme methionine synthase converts the substrate homocysteine to methionine through transfer of a methyl group from the co-substrate 5-methyltetrahydrofolate, while using vitamin B₁₂ as a cofactor (Doi 1989; Green 2011; Sauer 1977). Methionine is then used in the synthesis of S-adenosylmethionine (SAM), a major cellular methyl-donor in over 100 methylation reactions, including the methylation of arsenic. Arsenic excretion in humans involves methylation reactions, whereby methyl groups are transferred from SAM to inorganic arsenic by the enzyme Arsenic(III)-SAM-methyltransferase (Dheeman 2014; Schlebusch 2015). Given the critical role of folate in methylation reactions, low folate status, through inadequate intake of folates or through genetic variation affecting folate-metabolizing enzymes, may impede arsenic methylation and excretion and thereby exacerbate arsenic toxicity. In contrast, folic acid supplementation has been shown to improve symptoms of chronic arsenic exposure, including arsenical skin lesion (Ghose 2014) and oxidative DNA damage (Guo 2015). Previous randomised controlled clinical trials in the adult population in Bangladesh (Gamble 2007; Peters 2015) demonstrated that supplementation with folic acid lowers blood arsenic concentrations and increases urinary excretion of arsenic in a dimethylated form, suggesting folic acid supplementation as a preventive and/or therapeutic strategy for arsenic toxicity and arsenic-induced illnesses. Furthermore, an ongoing randomised controlled trial examines the arsenic-lowering effect of folic acid supplementation in arsenic-exposed Chinese populations (Xiao 2014). Despite the evidence from observational studies and clinical trials, no systematic reviews have been conducted to estimate the effects of folic acid supplementation on arsenic toxicity in children and in adult men and women.

Arsenic exposure may also affect folate metabolism. A metabolomic study in an animal
model reported a 9.9-fold decrease in urinary 5-methyltetrahydrofolate after four weeks exposure to arsenic (Lu 2014), suggesting that arsenic exhausts folate supply. A case-control study conducted in an arsenic-exposed population in rural Bangladesh also showed that arsenic exposure may reduce the effect of folic acid in the prevention of a specific form of neural tube defects, myelomeningocele (Mazumdar 2015a).

Inter-individual differences caused by genetic variations can affect the capacity to methylate arsenic (Gribble 2015; Schlebusch 2015). Genetic variants in folate-metabolizing enzymes have been shown to affect arsenic metabolism and retention in human and animal studies (Mazumdar 2015b; Schlawicke-Engstrom 2009; Wlodarczyk 2012; Wlodarczyk 2014). For example, a common genetic variant in the folate-metabolizing enzyme methylenetetrahydrofolate reductase (MTHFR), MTHFR C677T (i.e., C to T substitution at nucleotide position 677), has been associated with lower blood folate concentrations (Bailey 1999; Stover 2011; Tsang 2015) and increased sensitivity to arsenic exposure in a mouse model (Wlodarczyk 2012; Wlodarczyk 2014).

**Why it is important to do this review**

Inorganic arsenic is a common environmental toxin (ATSDR 2007; National Toxicology Program 2014), and an important public health burden given that there are 140 million people world-wide who are exposed through contaminated drinking water and food. Chronic arsenic exposure has been associated with neurotoxicity as well as an increased risk of cancer, congenital anomalies and skin diseases. Observational epidemiological studies support an adverse effect of arsenic on neurodevelopment in children (Parvez 2011; Tolins 2014; Wasserman 2007; Wasserman 2014; Wasserman 2016; Yorifuji 2016). Folate, which plays an
essential role in methylation reactions, may lower blood arsenic concentrations in arsenic-
exposed populations, thereby contributing to the prevention of arsenic-associated illnesses, including neurotoxicity in children (Hall 2009). The provision of folic acid through supplementation or fortification, or both, to reduce neural tube defects may have the added benefit of facilitating arsenic methylation and excretion to decrease arsenic toxicity, particularly in populations with folate deficiency (Dubey 2007; Gamble 2005; Gamble 2006; Gamble 2007; Pilsner 2009). This review is important because it will be the first to evaluate and summarize the evidence on folic acid supplementation or fortification, or both, for reducing arsenic toxicity and arsenic-related health outcomes in children and in adult men and women. The findings of this review will help inform future research and public policy, especially in arsenic-exposed populations.

OBJECTIVES

To assess the effects of provision of folic acid (through folic acid fortified foods or supplements), alone or in combination with other nutrients, on arsenic toxicity and arsenic-related health outcomes (i.e. neurocognitive function, skin lesions, congenital disorder and cancer) in arsenic-exposed populations.

METHODS

Criteria for considering studies for this review

Types of studies

We will only include the following study designs:

- randomised controlled trials (RCTs), with randomisation at either the individual- or
cluster-level; and

• quasi-RCTs (where allocation of treatment has been made, for example, by alternate allocation, date of birth or alphabetical order).

Types of participants
Arsenic-exposed populations of all ages and gender groups (including pregnant women) from any country.

We will define arsenic exposure through exposure to drinking-water arsenic levels above the WHO guideline value (i.e. 10 µg arsenic per litre water (WHO 2008)) or above the national permissible limits, or through exposure to arsenic levels in food (e.g., rice) that may results in measurable blood and urine arsenic exposure, based on the values defined by trial authors.

Types of interventions
Provision of folic acid (through fortified foods or supplements) alone or in combination with other nutrients.

Interventions may be given at any dose and for any duration, regardless of frequency of administration. A supplement may be in a tablet or capsule. Any form of folic acid fortification may be included. We will include studies with co-interventions (e.g., provision of arsenic water filtration systems, education, etc.), but only if the co-interventions were the same across study arms.

We plan to assess the following comparisons:

• folic acid supplements alone versus no intervention/placebo;

• folic acid supplements in combination with other nutrient supplements versus only other
nutrient supplements without folic acid (exact same formulation of other nutrients in both arms);  
- food fortified with folic acid alone versus unfortified food; and  
- food fortified with folic acid in combination with other nutrients versus food fortified with the exact same formulation of other nutrients (but no folic acid).

**Types of outcome measures**

We will consider both statistically and clinically significant minimal differences for each outcome listed below. We will collect outcomes measured at any time point post-intervention.

**Primary outcomes**

The primary outcomes to be considered across populations include:

- Any type of cancer (as defined by trial authors).
- All-cause mortality (as defined by trial authors).
- Neurocognitive function (as defined by trial authors, for example, including, but not limited to, the assessments of memory, attention, intelligence and other cognitive domains) (only in children).
- Any congenital anomalies (as defined by trial authors, including, but not limited to, neural tube defects, cleft lip and cleft palate, and detected, for example, by periconception and neonatal screenings) (only in pregnant women)

**Secondary outcomes**

- Blood or urinary arsenic concentration (μg/L)* (as measured by trial authors, for example, including, but not limited to, the measurements using the atomic absorption spectrometry and ICP-mass spectrometry; various forms of arsenic, for example,
monomethyl-, dimethyl-, or trimethyl-arsenic forms, will be considered).

- Blood folate concentration (nmol/L)* (as measured by trial authors, for example, including, but not limited to, the measurements using the mass spectrometry, competitive protein binding assays and microbiological assays).

- Serum/plasma homocysteine concentration (μmol/L)* (as measured by trial authors, for example, including, but not limited to, the measurements using high-performance liquid chromatography, gas chromatography-mass spectrometry and immunoassays).

- Skin lesions (as defined by trial authors, for example, including, but not limited to, the assessments of hyper- or hypo-pigmentation, keratosis and exfoliative dermatitis).

**Search methods for identification of studies**

We will search the electronic databases listed below for all available years without language restrictions. We will seek translations of documents, where needed. If we are unable to secure a translation, we will contact the editorial office of the Cochrane Developmental, Psychosocial and Learning Problems Group for support.

**Electronic searches**

We will search the following international and regional sources.

**International databases**

- Cochrane Central Register of Controlled Trials (CENTRAL; current issue) in the Cochrane Library, which includes the Cochrane Developmental, Psychosocial and Learning Problems Specialised Register.

- MEDLINE Ovid (1946 to current).

- MEDLINE In-Process & Other Non-Indexed Citations Ovid (current issue).
• MEDLINE Epub Ahead of Print Ovid (current issue).
• Embase Ovid (1980 to current).
• Social Science Citation Index Web of Science (1970 to current).
• Social Sciences Citation Index Web of Science (1970 to current).
• Conference Proceedings Citation Index - Science Web of Science (1990 to current).
• Conference Proceedings Citation Index - Social Science & Humanities Web of Science (1990 to current).
• Cochrane Database of Systematic Reviews (CDSR; current issue) part of the Cochrane Library.
• Database of Abstracts of Reviews of Effects (DARE; current issue) part of the Cochrane Library.
• CINAHL EBSCOhost (Cumulative Index to Nursing and Allied Health Literature; 1982 to current).
• POPLINE (www.popline.org).
• ClinicalTrials.gov (clinicaltrials.gov).
• WHO International Clinical Trials Registry Platform (ICTRP; apps.who.int/trialsearch).

Regional databases
3. Index Medicus for the South-East Asia Region (IMSEAR, imsear.hellis.org).
We will search MEDLINE using the search strategy in Appendix 1, which we will adapt for other databases, as appropriate. We will not limit the search by date or language.

**Searching other resources**

We will scan the reference lists of the included studies and relevant reviews to identify additional eligible studies. We will contact relevant study authors and experts in the field to identify any ongoing or unpublished studies.

**Data collection and analysis**

**Selection of studies**

After retrieving all articles from the searches, we will remove duplicate records. Two review authors (SB and EK) will independently screen the titles and abstracts of articles to assess eligibility based on the aforementioned criteria (see Criteria for considering studies for this review). When a title or abstract cannot be rejected with certainty, SB and EK will obtain the full-text report and independently screen it for a final assessment of inclusion. If we are unable to obtain full-text reports, we will request a copy from the trial authors. Disagreements that occur at any stage of the eligibility assessment process will be resolved through discussion with other authors, who, if necessary, will independently check the included and excluded studies.
Data extraction and management

Two review authors (SB and EK) will independently extract data from the included studies and record them on a piloted data extraction form designed for this review. The data extraction form will include the following information, and will be modified, if necessary.

- General information: title, authors, publication type (e.g., journal article, abstract, book chapter, etc.), country of study, funding source of study, year of study, and authors’ conflicts of interest.
- Details of study: study aim, design, inclusion/exclusion criteria, unit and method of randomisation, study location and duration, sample size, characteristics of participants, procedures for recruiting/selecting participants, method of allocation, and participant attrition.
- Intervention and control (comparison group): method used for implementation of intervention, duration/dose/frequency of intervention and control, type of intervention and control, co-intervention, number of participants allocated to intervention and control groups, and compliance.
- Outcomes: any measures of primary and secondary outcomes, time-points and method of outcome assessments, and blinding of outcome assessment.

Any disagreements between SB and EK during the data extraction process will be resolved through discussion with other authors. We will enter all extracted data into the latest version of Cochrane Review Manager software (Review Manager 2014), and we will check the data for accuracy. When the information regarding any of the above is insufficient or unclear, we will contact the trial authors to request further details of the study.
Assessment of risk of bias in included studies

Two review authors (SB and EK) will independently assess the risk of bias for each included study, based on the standard Cochrane ‘Risk of bias’ tool (Higgins 2011a). The authors will assign one of three ratings (low risk of bias; high risk of bias; or unclear risk of bias) to each of the domains listed below, with justifications for their judgements. Any disagreements will be resolved through discussion with other authors. When study information is insufficient or unclear, we will contact the trial authors to request further details. Examples of domain specific criteria for judgements of low, high or unclear risk of bias are shown in Table SA.1. More detailed criteria are provided in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011a, Table 8.5d).

Random sequence generation (checking for selection bias)

We will describe the method used to generate the allocation sequence and assess whether the sequence generation was suitable to minimize selection bias.

Allocation concealment (checking for selection bias)

We will describe the method used to conceal the allocation sequence and assess whether intervention allocation could have been foreseen in advance of, or during, enrolment.

Blinding of participants and personnel (checking for performance bias)

We will assess whether the study participants and personnel were blinded from knowledge of which intervention a participant received.

Blinding of outcome assessment (checking for detection bias)

We will assess whether the outcome assessors were blinded from knowledge of which
intervention a participant received.

*Incomplete outcome data (checking for attrition bias)*

We will assess whether the participants’ outcome data are missing due to attrition during the study or exclusions from the analysis. We will examine the reasons for attrition and exclusions, if reported, and whether the participants included in the analysis are exactly those who were randomised to the intervention and control groups.

*Selective reporting (checking for reporting bias)*

We will assess whether the included study reports only a subset of outcomes or only selective data for an outcome.

*Other sources of bias (checking for bias due to problems not covered by the domains above)*

We will assess other possible sources of bias, if any. For example, we will assess: whether the study has been claimed to be fraudulent; whether the study has contamination bias, which occurs when participants of the 'control' group inadvertently receive the treatment or are exposed to the intervention; and whether there are any other sources of bias not addressed in the other domains.

<table>
<thead>
<tr>
<th>TABLES SA.1. Examples of domain specific criteria for judgments of low, high or unclear risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Random sequence generation</strong></td>
</tr>
<tr>
<td><strong>Low risk of bias</strong></td>
</tr>
<tr>
<td><strong>High risk of bias</strong></td>
</tr>
<tr>
<td><strong>Unclear</strong></td>
</tr>
<tr>
<td>Risk of bias</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td><strong>Allocation concealment</strong></td>
</tr>
<tr>
<td>Low risk of bias</td>
</tr>
<tr>
<td>High risk of bias</td>
</tr>
<tr>
<td>Unclear risk of bias</td>
</tr>
<tr>
<td><strong>Blinding of participants and personnel</strong></td>
</tr>
<tr>
<td>Low risk of bias</td>
</tr>
<tr>
<td>High risk of bias</td>
</tr>
<tr>
<td>Unclear risk of bias</td>
</tr>
<tr>
<td><strong>Blinding of outcome assessment</strong></td>
</tr>
<tr>
<td>Low risk of bias</td>
</tr>
<tr>
<td>High risk of bias</td>
</tr>
<tr>
<td>Unclear risk of bias</td>
</tr>
<tr>
<td><strong>Incomplete outcome data</strong></td>
</tr>
<tr>
<td>Low risk of bias</td>
</tr>
<tr>
<td>High risk of bias</td>
</tr>
<tr>
<td>Unclear risk of bias</td>
</tr>
<tr>
<td><strong>Selective reporting</strong></td>
</tr>
<tr>
<td>Low risk of bias</td>
</tr>
<tr>
<td>High risk of bias</td>
</tr>
</tbody>
</table>
outcomes were reported incompletely and cannot be used.

<table>
<thead>
<tr>
<th>Unclear risk of bias</th>
<th>Unclear risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is no or insufficient information to permit a judgement of high or low risk of bias.</td>
<td>There is no or insufficient information to permit a judgement of high or low risk of bias.</td>
</tr>
</tbody>
</table>

Other sources of bias

<table>
<thead>
<tr>
<th>Low risk of bias</th>
<th>Baseline characteristics (related to outcome measures) or baseline outcome measures are similar across groups; an appropriate adjusted analysis was performed to account for differences in baseline measures across groups; or there are no other sources of bias.</th>
</tr>
</thead>
<tbody>
<tr>
<td>High risk of bias</td>
<td>Baseline characteristics (related to outcome measures) or baseline outcome measures are not similar across groups and no, or an inappropriate, adjusted analysis was performed; there is a contamination issue, whereby the experimental and control interventions get mixed; the study has been claimed to be fraudulent; or there are any other sources of bias.</td>
</tr>
</tbody>
</table>

| Unclear risk of bias                  | There is no or insufficient information to permit a judgement of high or low risk of bias. |

Measures of treatment effect

We will report treatment effects separately based on the types of outcome data (dichotomous or continuous). If there are too few trials (< two) or we are unable to combine trials, we will provide a narrative description of the results.

Dichotomous outcomes

We will present dichotomous outcome data as risk ratios (RR) with 95% confidence intervals (CIs). We will also show absolute measures of effect.

Continuous outcomes

We will use mean differences (MDs) with 95% CIs if continuous outcomes are measured on the same scale between trials. We will use standardized mean differences (SMD) with 95% CIs to combine trials that measure the same outcome but using different measurement methods. Where some studies have reported endpoint data and others have reported changes from baseline data (with errors), we will first divide them in subgroups; the subgroup with studies that used the
same scale but have reported a mixture of endpoint data and changes from baseline data will be combined in the meta-analysis. On the other hand, the subgroup with studies that used different scale but have reported a mixture of endpoint data and changes from baseline will not be combined, and they will make two more subgroups: the studies with only endpoint data and the studies with only changes from baseline data.

Unit of analysis issues

Cluster-randomised trials

We will consider combining the results from cluster- and individually-randomised trials if there is little heterogeneity between the studies, and if there is unlikely to be an interaction between the intervention effect and the choice of randomisation unit. We will conduct a sensitivity analysis to examine the impact of the randomisation unit (Sensitivity analysis). For cluster-randomised trials, we will examine whether the authors’ accounted for clustering in their analyses. If not, we will attempt to calculate effective sample sizes by using the intra cluster correlation coefficient (ICC), according to the formula provided in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011b). If neither the ICC nor the design effect is available from the study publication, we will contact the authors for further details or obtain external estimates of the ICC from similar studies. We will conduct a sensitivity analysis to examine the effect of variation in the ICC, where appropriate (Sensitivity analysis).

Studies with multiple treatment groups

For trials with more than two intervention groups, we will follow the approaches to avoid double counting of participants, as recommended by the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011b). We will present all intervention groups of a multi-arm study in
the ‘Characteristics of included studies’ tables. For the meta-analysis, we will assess which intervention group(s) of a multi-arm study is (are) relevant to the review based on the aforementioned criteria (see Criteria for considering studies for this review), and only include data from the directly relevant group(s). If a single study has more than one relevant intervention group or more than one control group, or both, we will create a single pair-wise comparison, where appropriate, by combining all relevant intervention groups into a single intervention group and all control groups into a single control group. For dichotomous outcomes, we will sum both the sample sizes and the number of people with events across groups. For continuous outcomes, we will combine means and standard deviations using the formula provided in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011b). In subgroup analyses, where the control group is shared by multiple intervention arms, we will divide the control group over the number of subgroup categories and include each pair-wise comparison separately (Subgroup analysis and investigation of heterogeneity). For dichotomous outcomes, we will divide the number of events and total participants. For continuous outcomes, we will divide the number of total participants only; the means and standard deviations will remain unchanged.

Cross-over trials

If we identify randomised trials with a cross-over design, we will consider using the data from the first period of the trial only, to avoid any potential risks of a carry-over effect.

Dealing with missing data

We will record attrition and missing outcome data for each included study on the data extraction form and report them in the ‘Risk of bias’ tables. When summary data for an outcome (e.g., standard deviations) are missing, we will base calculations on other reported measurements, if
possible. When pre-specified or expected outcome data are missing, we will contact the authors to request them. Where missing data are not supplied, we will report the available data alone without data imputation. For all relevant outcomes, we will attempt to conduct an intention-to-treat (ITT) analysis by including all participants randomised to each group irrespective of whether they actually received the allocated intervention. If a study report provides the outcome data only for the participants who completed the trial, or who complied with their allocated intervention, we will contact the authors to request the additional information needed to perform an ITT analysis. If no further information is available, we will conduct an available case analysis using data from participants whose results are known. We will conduct a sensitivity analysis to examine the impact of studies with missing data in the overall assessment of intervention effect (Sensitivity analysis), and describe the extent to which the missing data might affect the results and conclusions of the review.

Assessment of heterogeneity
We will assess clinical and methodological heterogeneity among studies by examining the variability in study design, participants, intervention, outcomes, and risk of bias. We will also assess statistical heterogeneity, which is likely to be a consequence of clinical or methodological heterogeneity, or both, by using the Chi² and P statistics included in the forest plots. We will consider a P value less than 0.10 in the Chi² test as evidence of heterogeneity of intervention effects. We will also use the I² statistic, which indicates the percentage of the variability due to heterogeneity rather than sampling error or chance, based on the thresholds listed below, as suggested in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011c). 0% to 40%: might not be important. 30% to 60%: may represent moderate heterogeneity. 50% to
90%: may represent substantial heterogeneity. 75% to 100%: considerable heterogeneity.

In addition, as an estimate of the between-study variability, we will report Tau2 from the random-effects meta-analysis. If there is heterogeneity among studies, we will explore the potential reasons for heterogeneity by conducting pre-specified subgroup analyses (Subgroup analysis and investigation of heterogeneity), and we will take caution in the interpretation of those results with high levels of unexplained heterogeneity.

**Assessment of reporting biases**

We will attempt to minimize reporting biases by comprehensively searching for eligible studies (including ongoing and unpublished studies) using multiple sources and databases. If there are more than 10 studies reporting the same outcome of interest, we will generate funnel plots and examine asymmetry that indicates the possibility of publication bias. We will also consider reasons for asymmetry other than publication bias such as differences in methodological quality among studies (Higgins 2011d). For example, smaller trials tend to use less rigorous methodological approaches than larger trials, which may result in spuriously larger intervention effects. If there is evidence of publication bias and small study effects, we will take it into account in the overall assessment and interpretation of intervention effects. We will also assess selective outcome reporting bias within each included study as described in the Assessment of risk of bias in included studies section. Where relevant outcome data are missing, we will contact the authors to request them. We will conduct a sensitivity analysis to explore the impact of studies with high levels of missing data in the overall assessment of results (Sensitivity analysis).

**Data synthesis**
We will conduct a meta-analysis to yield an overall (pooled) estimate of the intervention effect when more than one study can be appropriately combined (e.g., studies examining the same intervention and outcomes with comparable methods and approaches in similar populations). For example, we will assess the appropriateness of combining studies by considering whether they: 1) used comparable measurement tools or scales for outcome assessment; and 2) examined the same intervention and outcomes in similar populations (e.g., age (adults versus children) and reproductive status (pregnancy versus non-pregnancy)), and consider with the comparability of timing of the outcome measurement (i.e., scale comparable, timing comparable). To examine heterogeneity among studies, we will conduct subgroup analyses (see Subgroup analysis and investigation of heterogeneity) based on pre-specified factors (e.g., folic acid alone versus folic acid plus other nutrients).

We will perform a random-effects analysis since we anticipate natural heterogeneity among studies in terms of study populations, comparisons and interventions (e.g., doses and durations of intervention). We will pool the outcome data using the inverse variance method. We will attempt to conduct an ITT analysis by including all participants randomised to each group, where appropriate. If we are unable to combine data using a meta-analysis, we will provide a narrative summary of the results from individual studies.

Summary of findings
For each individual outcome, two review authors (SB and EK) will independently assess the quality of evidence across studies using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach (Balshem 2011). Disagreements will be resolved through discussion with other authors, who, if necessary, will independently conduct the quality assessment. We will grade evidence as one of four levels of quality (high, moderate,
low, or very low), depending on the presence of the following five factors: 1) within-study risk of bias (e.g., limitations in study design and implementation such as a lack of allocation concealment or blinding and a large loss to follow-up); 2) indirectness of evidence (e.g., indirect comparisons between interventions and the use of surrogate outcomes instead of health outcomes); 3) unexplained heterogeneity or inconsistency of results; 4) imprecision of results as indicated by wide CIs of effect estimates; and 5) high probability of publication bias. We will use GRADEprofiler: Guideline Development Tool (GRADEpro GDT 2015), to create a ‘Summary of findings’ table. The table will include all of the primary outcomes and the secondary outcomes indicated by an asterisk (*) (see the Types of outcome measures section) for each comparison listed in the Types of interventions.

Subgroup analysis and investigation of heterogeneity

Where sufficient data are available, we will conduct the subgroup analyses listed below, to investigate heterogeneity observed among studies. We will also examine whether the summary effects vary in relation to particular participant groups or types of intervention.

- Baseline folate status: deficient (defined as RBC folate <340 nmol/L or plasma/serum folate <10 nmol/L) versus non-deficient versus mixed/unknown/unreported.
- MTHFR C677T genotype: CC versus CT versus TT versus unknown/unreported, where power allows; or CC versus CT/TT versus unknown/unreported.
- Sex: male versus female versus mixed/unknown/unreported.
- Pregnancy: yes versus no versus mixed/unknown/unreported.
- Daily dose: less or equal than 400 μg/d folic acid versus greater than 400 μg/d folic acid.
- Nutrient composition: folic acid alone versus folic acid plus other nutrients.
• Duration: less than three months versus three months or more.

We will conduct subgroup analyses using Review Manager 2014. We will assess the differences among subgroups by inspecting the CIs of the summary estimates. Non-overlap of the CIs indicates a statistically significant difference in a treatment effect among the subgroups. We will also report the $I^2$ statistic for the interaction tests. Given that subgroup analyses are exploratory by definition, we will apply caution in the interpretation of the results of pre-planned subgroup analyses.

**Sensitivity analysis**

We will assess the robustness of the results and conclusions of the review by conducting sensitivity analyses to examine:

- the effect of excluding studies at high or unclear risk of bias in both the random sequence generation and allocation concealment domains and in one more of the other domains;
- the effect of excluding cluster-randomised trials where appropriate adjustments for clustering have not been made;
- the effect of different ICC values of cluster-randomised trials; and
- the effect of excluding studies with high levels (>20%) of missing outcome data.

**OTHER REFERENCES**

Ahmad 2001


Ahmed 2006

Almela 2002

Argos 2010

ATSDR 2007

Atta 2016

Axelson 1980

Bailey 1999
Bailey 2015


Balshem 2011


Banerjee 2013


Bearer 1995

Bearer CF. How are children different from adults? Environmental Health Perspectives 1995;103 Suppl 6:7-12.

Beaudin 2009


Botto 1999


Brandon 2014

Brandon EF, Janssen PJ, de Wit-Bos L. Arsenic: bioaccessibility from seaweed and rice, dietary

Carignan 2016


Carlin 2016


Caspi 2016


Chaineau 1990


Chen 2009


Chiou 2001

Choi 2010


Chávez-Capilla 2016


Cohen 2013


Concha 1998


Cordero 2015

Davis 2012

De-Regil 2015

deCastro 2014

DeSesso 2001

Dheeman 2014

Doi 1989

Dong 2009

167

Dubey 2007


EFSA 2009

European Food Safety Authority. Scientific opinion on arsenic in food. EFSA Journal 2009;7(10):1351 [199 pp.].

EFSA 2014


Ettinger 2009


Farzan 2013


FFI 2016


Fox 2008

Fox JT, Stover PJ. Folate-mediated one-carbon metabolism. Vitamins and Hormones 2008;79:1-
Fängström 2009

Gamble 2005

Gamble 2006

Gamble 2007

Gardner 2011

Gardner 2012
methyltransferase genotype independently influence the arsenic methylation phenotype. Pharmacogenet Genomics 2012;22(7):508-16.

Ghose 2014


Ghosh 2013


Gilbert-Diamond 2011


GRADEpro GDT 2015


Green 2011


Greene 2014


Gribble 2013

Gribble MO, Crainiceanu CM, Howard BV, Umans JG, Francesconi KA, Goessler W, et al.

Gribble 2015


Grosse 2016


Guo 2015


Hall 2006


Hall 2009


Hamadani 2010

Hamadani 2011


Han 2011


Higgins 2011a


Higgins 2011b


Higgins 2011c


Higgins 2011d

Hill 2008

Hill 2009

Hojsak 2015

Hopenhayn-Rich 2000
Hsieh 2014

Huyck 2007

IARC 1980

IARC 2012

IPCS 2001

Jansen 2016
Karagas 2016

Khan 2015

Kile 2016

Kim 2011

Kippler 2016

Kurzzius-Spencer 2014
Kwok 2006


Laine 2015


Li 2011


Lindberg 2008


Ljung 2011


Lu 2014


Lynch 2014


Ma 2008


Mazumdar 2015a


Mazumdar 2015b


Mazumder 2010

Meharg 2008

Melkonian 2013

Milton 2005

Moon 2013

Morrissey 1983

National Toxicology Program 2014

Naujokas 2013

Navas-Acien 2008


Nielsen 2010


O'Bryant 2011


Parvez 2011


Peters 2015


Pfeiffer 2015

Pilsner 2009

Pinyayev 2011

Prakash 2016

Rader 2006

Rahman 2007

Review Manager 2014

Rintala 2014


Rodriguez-Barranco 2016


Rodriguez-Lado 2013


Rosado 2007


Rose 2007


Rubin 2014


Sanders 2012

Sauer 1977


Schlawicke-Engstrom 2009


Schlebusch 2015


Seo 2016


Sidhu 2015


Signes-Pastor 2016

Skröder Löveborn 2016

Smith 2000

Smith 2002

Smith 2006

Sohel 2009

Sohn 2014

Stone 2008

Stover 2004

Stover 2011


Styblo 2000


Tolins 2014


Tsai 2003


Tsang 2015


Tseng 2005

Tseng CH, Huang YK, Huang YL, Chung CJ, Yang MH, Chen CJ, et al. Arsenic exposure,

Tsuji 2015


USPSTF 2009


Vahidnia 2007


Vahter 2000


Valenzuela 2005


von Ehrenstein 2007

Wasserman 2007

Wasserman 2014

Wasserman 2016

WHO 2001

WHO 2008

Wlodarczyk 2001
Wlodarczyk 2012

Wlodarczyk 2014

Wu 1989

Wu 2011

Xiao 2014

Yorifuji 2016

Yu 2007

Yuan 2007


Yuan 2010


Zavala 2008a


Zavala 2008b


Zhang 2014

APPENDIX 1

1 MEDLINE Search Strategy (Ovid)

1 Metals, Heavy/

2 Heavy Metal Poisoning, Nervous System/

3 (metalloid$ or metal poison$ or heavy metal$).tw,kw.

4 Arsenic/

5 Arsenic Poisoning/

6 exp Arsenicals/

7 arsen$.mp.

8 (arsenic$ or arsenite$ or arsenate$).mp.

9 (monomethylarsonic or mono-methylarsonic).mp.

10 (monomethylarsonous or mono-methylarsonous).mp.

11 (dimethylarsinous or di-methylarsinous).mp.

12 (trimethylarsine or tri-methylarsine).mp.

13 or/1-12

14 exp Folic Acid/

15 Folic Acid Deficiency/

16 folate$.mp.

17 folic.mp.

18 folinic$.mp.

19 folacin$.mp.

20 vitamin B9.mp.

21 5-methyltetrahydrofolate.mp.
22 5-methylTHF.mp.
23 metafolin.mp.
24 leucovorin.mp.
25 Micronutrients/
26 Food, Fortified/
27 Functional Food/
28 (food$ adj3 (fortif$ or functional$ or supplement$1)).tw,kw.
29 (micro-nutrient$ or micronutrient$).tw,kw.
30 or/14-29
31 13 and 30
APPENDIX B

Copy right approval forms
Permissions

3/31/2017

Sajin Bae
210 Savage Hall
Cornell University
Ithaca NY 14853
sb749@cornell.edu

Dear Sir or Madame,

We are in receipt of your request to reproduce your Open Access article for use in your dissertation

Impact of folic acid fortification on global DNA methylation and one-carbon biomarkers in the Women's Health Initiative Observational Study Cohort
Epigenetics 9 (3): 396-403.
DOI: 10.4161/epi.27323

You may use the published version of your article.

This permission is all for print and electronic editions.

We will be pleased to grant you permission free of charge on the condition that:

This permission is for non-exclusive English world rights. This permission does not cover any third party copyrighted work which may appear in the material requested.

Full acknowledgment must be included showing authors, year published, article title, and full Journal title, reprinted by permission of Taylor & Francis LLC (http://www.tandfonline.com).

Thank you very much for your interest in Taylor & Francis publications. Should you have any questions or require further assistance, please feel free to contact me directly.

Sincerely,

Mary Ann Muller
Permissions Coordinator
Telephone: 215.606.4334
E-mail: maryann.muller@taylorandfrancis.com
This Agreement between Sajin Bae ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number: 4100821476017
License date: May 02, 2017
Licensed Content Publisher: John Wiley and Sons
Licensed Content Publication: Cochrane Database of Systematic Reviews
Licensed Content Title: Provision of folic acid for reducing arsenic toxicity in arsenic-exposed children and adults
Licensed Content Author: Sajin Bae, Elena Kamynina, Adetutu F. Farinola, Marie A. Caudill, Patrick J. Stover, Patricia A. Cassano, Robert Berry, Juan Pablo Peña-Rosas
Licensed Content Date: May 2, 2017
Licensed Content Pages: 1
Type of use: Dissertation/Thesis
Requestor type: Author of this Wiley article
Format: Print and electronic
Portion: Full article
Will you be translating?: No
Title of your thesis / dissertation: Influence of B-vitamins on one-carbon metabolism and associations with cancer risk and reproductive state
Expected completion date: May 2017
Expected size (number of pages): 205
Requestor Location: Sajin Bae
Cornell University

ITHACA, NY 14850
United States
Attn: Sajin Bae

Publisher Tax ID: EU826007151
Billing Type: Invoice
Billing Address: Sajin Bae
Cornell University

ITHACA, NY 14850
United States
Attn: Sajin Bae

Total: 0.00 USD
TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.

- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, and any CONTENT (PDF or image file) purchased as part of your order, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor
formatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND
WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a
party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified
mail, return receipt requested, at the last known address of such party.

**WILEY OPEN ACCESS TERMS AND CONDITIONS**

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

**The Creative Commons Attribution License**

The [Creative Commons Attribution License (CC-BY)](https://creativecommons.org/licenses/by/) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

**Creative Commons Attribution Non-Commercial License**

The [Creative Commons Attribution Non-Commercial (CC-BY-NC)](https://creativecommons.org/licenses/by-nc/)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

**Creative Commons Attribution-Non-Commercial-NoDerivs License**

The [Creative Commons Attribution-Non-Commercial-NoDerivs License (CC-BY-NC-ND)](https://creativecommons.org/licenses/by-nc-nd/) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

**Use by commercial "for-profit" organizations**

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library

http://olabout.wiley.com/WileyCDA/Section/id-410895.html

**Other Terms and Conditions:**

v1.10 Last updated September 2015

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or