Chronic arsenic exposure in Bangladesh and the United States: from nutritional influences on arsenic methylation to arsenic-induced epigenetic dysregulation

Anne Kristina Bozack

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

ABSTRACT

Chronic arsenic exposure in Bangladesh and the United States: from nutritional influences on arsenic methylation to arsenic-induced epigenetic dysregulation

Anne Kristina Bozack

Background: Chronic arsenic (As) exposure in a global public health concern. Arsenic exposure through drinking water affects over 140 million people in at least 70 countries, including 40 million people in Bangladesh. In the United States (US), 2.4 million people rely on private wells or public water systems with As levels above the US maximum contaminant level.

Ingested inorganic arsenic (InAs) is methylated to monomethyl (MMAs)- and dimethyl (DMAs)-arsenical species using the methyl donor *S*-adenosylmethionine (SAM). Full methylation of InAs to DMAs decreases As toxicity and facilitates urinary As excretion. Arsenic methylation capacity is influenced by nutrients involved in one-carbon metabolism (OCM), the biochemical pathway that synthesizes SAM. Folate recruits one-carbon units for the remethylation of homocysteine and the synthesis of SAM. The availability of one-carbon units is also impacted by nutrients including the alternative methyl donor betaine, its precursor choline, and possibly the cofactor vitamin B_{12}. In addition, As methylation capacity may also be influenced by creatine; an estimated 50% of SAM is consumed by the final step of endogenous creatine synthesis.

The adverse health outcomes associated with chronic As exposure include impaired intellectual function, cardiovascular disease, diabetes, inflammation, and cancers of the bladder, lung, kidney, liver, and skin. *In utero* As exposure is associated with adverse birth outcomes include decreased birth weight and gestational age. Elevated health

risks persist after exposure has been reduced or ended, leading to the hypothesis that epigenetic dysregulation, including changes in DNA methylation, may be a biological mechanism linking As exposure to health outcomes.

Objectives: This research has three main objectives: (1) to investigate the influence of OCM nutritional factors on As methylation by evaluating effects of folic acid (FA) and creatine supplementation on As methylation capacity, and effect modification by baseline status of OCM-related nutrients; (2) to examine associations between As exposure and loci-specific DNA methylation in an epigenome-wide association study (EWAS); and (3) to assess mediation of the association between *in utero* As exposure and birth outcomes (i.e., gestational age and birth weight) by DNA methylation of target genes identified in an EWAS, as well as the candidate gene DNA methyltransferase 3 alpha (*DNMT3A*), a protein-coding gene involved in *de novo* DNA methylation.

Methods: This research used data from three studies of As-exposed individuals. To address the first objective, we used data from the Folic Acid and Creatine Trial (FACT), a 24-week randomized clinical trial of FA (400 or 800 μg/day) and/or creatine supplementation (3 g/day or 3 g creatine and 400 μg FA/day) among As-exposed adults in Bangladesh recruited independent of folate status (N = 622). We investigated overall FA and creatine treatment effects on mean within-person changes in As metabolite proportions in urine compared to the placebo group (weeks 0 to 12). Rebound of As methylation capacity following the cessation of FA supplementation was assessed from weeks 12 to 24. We also assessed effect modification by baseline choline, betaine, vitamin B_{12}, and plasma folate of treatment effects on changes in homocysteine, guanidinoacetate (GAA) (biomarkers of OCM and endogenous creatine synthesis, respectively), total blood As, and urinary As metabolite proportions and indices.

To address the second objective, we used data from the Strong Heart Study (SHS), a population-based prospective cohort of American Indians with low-moderate levels of

As exposure. DNA methylation was measured in 2,325 participants using the Illumina MethylationEPIC array, which interrogates > 850,000 loci. We tested for differentially methylated positions (DMPs) and regions (DMRs), and conducted gene ontology (GO) enrichment analysis to understand functions of genes containing differential methylation.

To address the third objective, we used data from a prospective birth cohort in Bangladesh. In a discovery phase, an EWAS was conducted to identify CpGs with methylation measured in cord blood that are associated with maternal water As levels and birth outcomes (N = 44). In a validation phase, DNA methylation in cord blood was measured using bisulfite pyrosequencing at three target CpGs annotated to *miR124-3*, *MCC*, and *GNAL* (N = 569). We applied structural equation models (SEMs) to assess mediation of the association between *in utero* As exposure and gestational age by DNA methylation. In addition, mediation of the association between *in utero* As exposure and birth outcomes by DNA methylation of the candidate gene DNA methyltransferase alpha (*DNMT3A*) was assessed.

Results: In FACT, the mean within-person decreases %InAs and %MMAs and increase in %DMAs were greater among all groups receiving FA supplementation at weeks 6 and 12 compared to placebo ($P < 0.05$) (**Chapter 3**). Stratified by median choline and betaine concentrations at baseline, we observed a trend towards greater FA treatment effects among participants with levels below the median of both nutrients compared to participants above the median (**Chapter 4**). Among participants who discontinued FA supplementation, at week 24, %InAs and %DMAs were not significantly different than baseline levels, suggesting a rebound in As methylation capacity with cessation of FA supplementation. We observed a significantly greater mean within-person decreases in %MMAs with creatine supplementation compared to placebo at weeks 1, 6, and 12; mean within-person changes in %InAs and %DMAs did not differ significantly

between the creatine and placebo groups (**Chapter 3**). The mean within-person decrease in urinary %MMAs at week 12 with creatine treatment was significantly greater than placebo among participants with baseline choline concentrations below the median, but did not differ from placebo among participants with choline concentrations above the median (**Chapter 4**).

In an EWAS conducted in SHS, we identified 20 DMPs associated with urinary As levels at $FDR < 0.05$; five DMPs were significant at $P_{Bonferroni} < 0.05$ (**Chapter 5**). The top significant CpG, cg06690548, was located in solute carrier family 7 member 11 (*SLC7A11*), part of the amino-acid transporter cystine:glutamate antiporter system x_c^-, which is involved in biosynthesis of the endogenous antioxidant glutathione (GSH). Additional Bonferroni-significant CpGs were located in *ANKS3*, *LINGO3*, *CSNK1D*, and *ADAMTSL4*. We identified one FDR-significant DMR (chr11:2,322,050-2,323,247) including the open reading frame C11orf21 and tetraspanin 32 (*TSPAN32*).

Mediation of the association between *in utero* As exposure and birth outcomes by cord blood DNA methylation was assessed in a Bangladeshi birth cohort. In the discovery phase (N = 44), the association between maternal water As levels and gestational age was fully mediated by DNA methylation of the top 10 CpGs associated with both variables. In a discovery phase (N = 569), there were significant indirect effects of maternal water As levels on gestational age through DNA methylation of *miR124-3* and *MCC*; the indirect effect through DNA methylation of *GNAL* was not significant (**Chapter 6**). In an adjusted SEM including *miR124-3* and *MCC*, mediation of the association between *in utero* As exposure and gestational age by DNA methylation of *miR124-3* was borderline significant ($P = 0.06$); DNA methylation of *MCC* did not act as a mediator. We also assessed mediation by DNA methylation of *DNMT3A* (**Chapter 7**). In an adjusted SEM including birth weight and gestational age, there was a significant indirect effect of maternal toenail As levels on gestational age through

DNMT3A methylation, the indirect effect on birth weight was borderline significant (P = 0.082). However, the indirect effects of maternal toenail As levels on birth weight through all pathways including gestational age were statistically significant. A doubling in maternal toenail As concentrations had a total effect of a decrease in gestational age of 2.1 days and a decrease in birth weight of 28.9 g.

Conclusions: Results from FACT (**Chapters 3 and 4**) provide evidence of the associations between OCM-related nutrients and As methylation capacity. Specifically, FA and creatine supplementation may increase As methylation capacity by increasing the availability of SAM, and treatment effects may be greater among individuals with low betaine and choline status, respectively. In addition, results reported in **Chapters 5-7** support the hypotheses that chronic As exposure is associated with epigenetic dysregulation, and that changes in the epigenome may mediate the association between As exposure and adverse health effects. Findings from the research presented here may help inform public health interventions to reduce the adverse health effects of chronic As exposure. However, further research is needed to fully understand the biological mechanism that influence As methylation and that underlie the associations between chronic As exposure and adverse health outcomes.

Contents

List of Figures

List of Tables

List of Abbreviations

400FA	400 µg folic acid per day
450K	Illumina Infinium HumanMethylation BeadChip
850K	Infinium MethylationEPIC BeadChip
5-caC	5-carboxylcytosine
5-fC	5-formylcytosine
5-hmC	5-hydroxymethylcytosine
5-mC	5-methylcytosine
5-methyl-THF	5-methyltetrahydrofolate
5,10-methylene-THF	5,10-methylene-tetrahydrofolate
8-oxodG	8-oxo-29-deoxyguanosine
800FA	800 µg folic acid per day
As	Arsenic
AS3MT	Arsenic (+3 oxidation state) methytransferase
bAs	Blood arsenic
BEAR	Biomarkers of Exposure to Arsenic Study
BEST	Bangladesh Vitamin E and Selenium Trial
BHMT	Betaine homocysteine methyltransferase
CC	Case-control
Chr	Chromosome
CK1	Casein kinase 1
CpG	Cytosine-guanine dinucleotide
CS	Cross-sectional
CSNK1D	Casein kinase 1 delta
DCH	Dhaka Community Hospital
DHF	Dihydrofolate
DMAs	Dimethyl-arsenical species
$DMAs^V$	Dimethylarsinic acid
DMP	Differentially methylated position
DMR	Differentially methylated region
DNAm	DNA methylation
DNMT	DNA methyltransferase
DNMT3A	DNA methyltransferase 3 alpha
DNMT3B	DNA methyltransferase 3 beta
dTMP	2-deoxy-thymidine-5-monophosphate

dUMP	2-deoxy-uridine-5-monophosphate
EPA	Environmental Protection Agency
EWAS	Epigenome-wide association study
FA	Folic acid
FACT	Folic Acid and Creatine Trial
FIML	Full information maximum likelihood approach
FOX	Folate and Oxidative Stress Study
FDR	False discovery rate
g	Grams
GAA	Guanidinoacetate
GAMT	Guanidinoacetate N-methyltransferase
GNAL	G protein subunit alpha L
GO	Gene ontology
GSH	Glutathione
h	hour
HEALS	Health Effects of Arsenic Longitudinal Study
HK1	Hexokinase 1
HUVEC	Human umbilical vein endothelial cells
InAs	Inorganic arsenic (InAs)
$InAs^{III}$	Arsenite
$InAs^{V}$	Arsenate
LINGO3	Leucine rich repeat and Ig domain containing 3
MCC	Mutated in colorectal cancers
MCL	Maximum contaminant level
miR123-4	microRNA 124-3
miRNAs	microRNAs
MMAs	Monomethyl-arsenical species
$MMAs^{III}$	Monomethylarsonous acid
$MMAs^{V}$	Monomethylarsonic acid
MRP	Multidrug resistance protein
MTR	Methionine synthase
NK	Naural killer cells
NTD	Neural tube defect
OCM	One-carbon metabolism
OR	Odds ratio
PARP-1	Poly(ADP-ribose)polymerase
PC	Phosphatidylcholine
PCA	Principal component analysis
PCFT	Proton coupled folate transporter
PMI	Primary methylation index
PLP	Pyridoxal phosphate
RBC	Red blood cell

RCT	Randomized controlled trial
ROS	Reactive oxygen species
SAH	*S*-adenosylhomocysteine
SAM	*S*-adenosylmethionine
SEM	Structural equation model
SHS	Strong Heart Study
SLC	Solute carrier
SLC19A1	Solute carrier family 19 (folate transporter), member 1
SLC7A11	Solute carrier family 7 member 11
SLC7A5	Solute carrier family 7 member 5
SMI	Secondary methylation index
SNP	Single nucleotide polymorphism
TDG	Thymine DNA glycosylase.
TET	Ten-eleven translocation
THF	Tetrahydrofolate
TMAO	trimethylamine oxide
TS	Thymidylate synthetase
TSPAN32	Tetraspanin 32
TSS	Transcription start site
TYMS	Thymidylate synthetase
UNICEF	United Nations Children's Fund
US	United States
UTR	Untranslated region
WHO	World Health Organization

Acknowledgements

First, I would like to thank the participants in Bangladesh and the United States who have been part of this research, and without whom this dissertation would not have been possible. Although the toxicity of arsenic has been known for millennia, millions of people throughout the world are still chronically exposed to arsenic. To gain a better understanding of the health effects of arsenic exposure, factors that increase susceptibility, and underlying biological mechanisms, we rely on the engagement and commitment of study participants.

This work would also not have been possible without the support of my mentors and committee members. Research, and this dissertation, is the result of collaboration. Thank you to my primary mentor, Dr. Mary Gamble, for giving me a place in your lab, and your patience over the last five years. Your ongoing support in research, training, and professional development has provided me with growth and opportunity.

I was lucky to have to opportunity to work closely with Drs. Ana Navas-Acien and Molly Kile on my dissertation research. I would like to thank Dr. Navas-Acien for giving me the opportunity to work with the Strong Heart Study. I would also like to thank Dr. Kile for serving as my mentor during my KC Donnelly Externship at Oregon State University. Not only did you provide me with research mentorship, but you also welcomed me into your lab and helped make Corvallis home for me during the program. My additional committee members, Drs. Andres Cardenas, Julie Herbstman, and Shuang Wang have also provided valuable guidance throughout my research and professional development.

I was also supported by members of the Columbia University Superfund Research

Program, including Dr. Joseph Graziano, Nancy Loiacono, Dr. Xinhua Liu, Vesna Ilievski, Vesna Slavkovich, Arce Domingo, lab members, and field staff in Bangladesh. Many other members of the Columbia community have supported me and helped me navigate this process, including Nina Kulacki, Fernando Luque.

Thank you to my fellow students at Columbia University, current and former, for your friendship, smiles, and laughter. This program would not have been the same without Dan Carrion, Whitney Cowell, Tiffany Sanchez, Ashlinn Quinn, Catilin Howe, Alex Heaney, Sebastian Roland, Carlos Gould, Lizzy Gibson, Roheeni Saxena, Annie Nigra, and Ahlam Abuawad. I addition, I would like to thank students at Oregon State University, including Barrett Welch and Holly Dixon, for welcoming me during my KC Donnelly Externship.

I am thankful to have received support as a trainee with the national NIEHS Superfund Research Program, including experience gained through a KC Donnelly Externship. I have also been fortunate to receive funding support through an NIEHS F31 fellowship, the Columbia University NIEHS Center for Environmental Health in Northern Manhattan, and additional funding from the Department of Environmental Health Sciences.

On a more personal note, I am grateful for my family and friends, including Lucia Macias, Carlos Salazar, Alena Alasheeva, Andrew Ressler, Phong Bui, Jyoti Gupta, Stephanie Myers, Sejal Vashi, Richard Remigio, and many others in Berkeley and New York throughout the years. Finally and most importantly, thank you to my parents for your support and for teaching me curiosity and adventure. I could not have come this far without your endless love and encouragement.

Chapter 1

Statement of hypotheses

1.1 Overview

Over 140 million people are exposed to arsenic (As) concentrations in drinking water greater than 10 μg/L, the World Health Organization (WHO) guideline and United States (US) Environmental Protection Agency (EPA) maximum contaminant level (US EPA Office of Water, 2001; World Health Organization, 2012), including 40 million people in Bangladesh (Bangladesh Bureau of Statistics and United Nation Children's Fund, 2015). Although As exposure is less common in the US, 296,000 people are served by public water systems with As contamination over 10 μg/L (U.S. Environmental Protection Agency, 2018) and 2.1 million individuals rely on private well water with elevated As levels (Ayotte et al., 2017).

Ingested inorganic arsenic (InAs) undergoes a series of reduction and oxidative methylation reactions. InAs is methylated to monomethyl (MMAs)- and dimethyl (DMAs)-arsenical species using the methyl donor *S*-adenosylmethionine (SAM) (Challenger, 1945). Full methylation of InAs to DMAs decreases As toxicity and facilitates urinary As excretion (Tice et al., 1997; Vahter and Marafante, 1987). Arsenic methylation capacity is influenced by nutrients involved in one-carbon metabolism (OCM) (Bozack et al., 2018), the biochemical pathway that synthesizes SAM.

Chronic As exposure is associated with numerous adverse health effects including impaired intellectual function, cardiovascular disease, diabetes, inflammation, and cancers of the bladder, lung, kidney, liver, and skin (IARC Working Group, 2009; Moon et al., 2017; National Research Council, 2013). Elevated risks persist after exposure has been reduced or ended (Smith et al., 2012; Steinmaus et al., 2013, 2014), suggesting epigenetic dysregulation, including changes in DNA methylation, may be a biological mechanism linking As exposure to health outcomes.

This dissertation will investigate the association between OCM nutrients and As methylation capacity, the effect of As exposure on epigenetic dysregulation, and

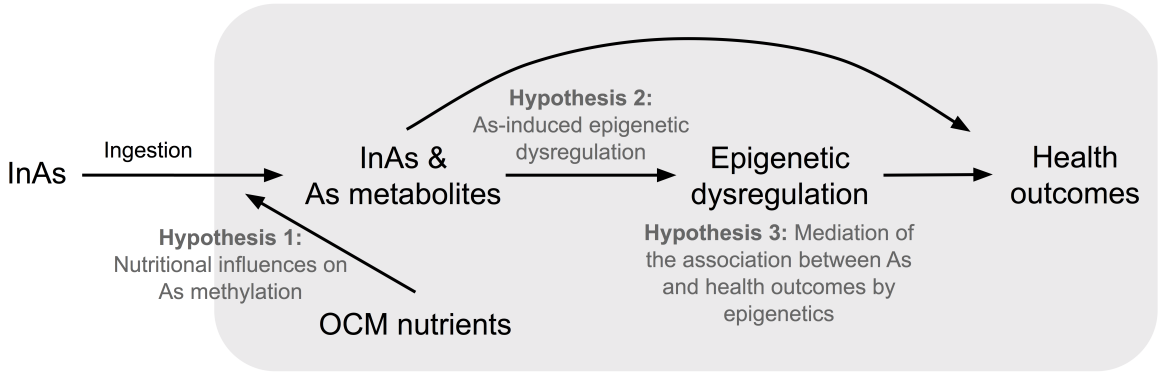


Figure 1.1: Overview of hypotheses.

the mediating role of epigenetics in the association between As exposure and health outcomes (Figure 1.1).

1.2 Hypothesis 1: Nutritional influences on arsenic methylation

Folate and creatine are two nutrients related to OCM. In the Folic Acid and Creatine Trial (FACT), a randomized controlled trial (RCT) among As-exposed adults in Bangladesh, our group previously observed a larger decrease in ln(blood As (bAs) concentrations) with 12 weeks of 800 μg folic acid (FA)/day supplementation compared to placebo (Peters et al., 2015), suggesting that FA supplementation increased As excretion my influencing As methylation capacity. We hypothesize that FA and creatine supplementation will increase As methylation as measured by the proportion of As metabolites in urine.

Hypothesis 1a, FA and creatine treatment effects of As methylation:

We will examine the treatment effects of 12 weeks of FA (400 or 800 μg/day), creatine (3 g/day), and creatine + FA supplementation (3 g creatine + 400 μg FA/day) on the mean within-person changes in %InAs, %MMAs, %DMAs measured in urine compared to placebo. It is expected that the mean within-person decreases in %InAs and %MMAs and increase in %DMAs with FA and create treatment will exceed that of placebo.

Hypothesis 1b, rebound of FA and creatine treatment effects:

We will evaluate changes in %InAs, %MMAs, and %DMAs in urine following cessation of FA supplementation (weeks 12 to 24). We expect that As methylation capacity will rebound after the cessation of FA supplementation.

Hypothesis 1c, effect modification by baseline nutritional status:

We will examine differences in treatment effects on changes in total homocysteine concentrations (a biomarker of OCM), guanidinoacetate concentrations (GAA, a biomarker of endogenous creatine biosynthesis), bAs concentrations, and As metabolite proportions and indices stratified by baseline median choline, betaine, vitamin B_{12}, and folate. The effect of FA and creatine supplementation is expected to be greater among participants with low status of other OCM nutrients.

Hypothesis 1 findings will be reported in **Chapters 3 and 4**.

1.3 Hypothesis 2: Epigenome-wide associations with arsenic exposure

Associations between chronic As exposure and locus-specific DNA methylation have been observed in previous epigenome-wide association studies (EWAS) of exposure *in utero* (Broberg et al., 2014; Cardenas et al., 2015; Gliga et al., 2018; Green et al., 2016; Kaushal et al., 2017; Kile et al., 2014; Rojas et al., 2015) and adulthood (Ameer et al., 2017; Argos et al., 2015; Demanelis et al., 2019; Guo et al., 2018). We will conduct an EWAS in the Strong Heart Study (SHS), a population-based prospective cohort of American Indians with low-moderate levels of As exposure. We hypothesize that As concentrations measured in urine will be associated with DNA methylation at individual loci and regions.

Hypothesis 2a, differentially methylated positions and regions:

We will examine the associations between urinary As levels and DNAm at individual loci and regions. In addition, associations between urinary As levels and DNA methylation at loci previously identified as differentially methylated with As exposure will be assessed.

Hypothesis 2b, gene ontology enrichment analysis:

We will conduct a gene ontology (GO) enrichment analysis to identify biological and molecular functions associated with genes containing CpGs that are differentially methylated by As exposure. We expect to identify GO terms with biological relevance to As exposure.

Hypothesis 2 findings will be reported in **Chapter 5**.

1.4 Hypothesis 3: *In utero* arsenic exposure and birth outcomes, mediation by DNA methylation

In utero As exposure has been associated with adverse birth outcomes including reductions in birth weight (reviewed in Bloom et al. (2014); Milton et al. (2017); Zhong et al. (2019)) and gestational age (Röllin et al., 2017; Xu et al., 2011). We hypothesize that epigenetic dysregulation mediates the association between *in utero* As exposure and birth outcomes. The relationships between maternal exposure to As during pregnancy, birth weight, gestational age, and DNA methylation will be investigated in a prospective birth cohort of As-exposed women in Bangladesh.

Hypothesis 3a, epigenome-wide approach:

CpGs associated with maternal water As concentrations and gestational age were identified using an epigenome wide approach. Top CpGs identified mediated the association between As exposure and gestational age. In a validation phase, using structural equation models (SEMs) we will evaluate mediation of the association between *in utero* As exposure and gestational age by methylation of three top CpGs measured by pyrosequencing.

Hypothesis 3b: DNAm *DNMT3A*:

DNA Methyltransferase 3 Alpha (DNMT3A) is expressed during embryogenesis and is involved in *de novo* DNA methylation. We selected *DNMT3A* as a candidate gene involved in mediation of the association between *in utero* As exposure and birth outcomes. Using SEMs, we will evaluate the relationships between *in utero* As expo-

sure measured by maternal toenail As concentrations, gestational age, birth weight, and *DNMT3A* methylation using pyrosequencing.

Hypothesis 3 findings will be reported in **Chapters 6 and 7**.

1.5 Bibliography

Ameer, S S; Engström, K; Hossain, M B; Concha, G; Vahter, M, and Broberg, M. Arsenic exposure from drinking water is associated with decreased gene expression and increased DNA methylation in peripheral blood. *Toxicology and Applied Pharmacology*, 321:57–66, 2017. doi: 10.1016/J.TAAP.2017.02.019.

Argos, M; Chen, L; Jasmine, F; Tong, L; Pierce, B L; Roy, S; Paul-Brutus, R; Gamble, M V; Harper, K N; Parvez, F; Rahman, M; Rakibuz-Zaman, M; Slavkovich, V; Baron, J A; Graziano, J H; Kibriya, M G, and Ahsan, H. Gene-specific differential DNA methylation and chronic arsenic exposure in an epigenome-wide association study of adults in Bangladesh. *Environmental Health Perspectives*, 123(1):64–71, 2015. doi: 10.1289/ehp.1307884.

Ayotte, J D; Medalie, L; Qi, S L; Backer, L C, and Nolan, B T. Estimating the high-arsenic domestic-well population in the conterminous United States. *Environmental Science and Technology*, 51(21):12443–12454, 2017. doi: 10.1021/acs.est.7b02881.

Bangladesh Bureau of Statistics and United Nation Children's Fund. Bangladesh multiple indicator cluster survey 2012-2013 final report. Technical report, Dhaka, Bangladesh, 2015.

Bloom, M S; Surdu, S; Neamtiu, I A, and Gurzau, E S. Maternal arsenic exposure and birth outcomes: a comprehensive review of the epidemiologic literature focused on drinking water HHS Public Access. *International Journal of Hygiene and Environmental Health*, 217(7):709–719, 2014. doi: 10.1016/j.ijheh.2014.03.004.

Bozack, A K; Saxena, R, and Gamble, M V. Nutritional influences on one-carbon metabolism: effects on arsenic methylation and toxicity. *Annual Review of Nutrition*, 38:401–429, 2018. doi: 10.1146/annurev-nutr-082117-051757.

Broberg, K; Ahmed, S; Engström, K; Hossain, M B; Jurkovic Mlakar, S; Bottai, M; Grandér, M; Raqib, R, and Vahter, M. Arsenic exposure in early pregnancy alters genome-wide DNA methylation in cord blood, particularly in boys. *Journal of Developmental Origins of Health and Disease*, 5(4):288–98, 2014. doi: 10.1017/S2040174414000221.

Cardenas, A; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Wright, R O; Christiani, D C, and Kile, M L. *In utero* arsenic exposure and epigenome-wide associations in placenta, umbilical artery, and human umbilical vein endothelial cells. *Epigenetics*, 10(11):1054–1063, 2015. doi: 10.1080/15592294.2015.1105424.

Challenger, F. Biological methylation. *Chemical Reviews*, 36(3):315–361, 1945. doi: 10.1021/cr60115a003.

Demanelis, K; Argos, M; Tong, L; Shinkle, J; Sabarinathan, M; Rakibuz-Zaman, M; Sarwar, G; Shahriar, H; Islam, T; Rahman, M; Yunus, M; Graziano, J H.; Broberg, K; Engström, K; Jasmine, F; Ahsan, H, and Pierce, B L. Association of arsenic exposure with whole blood DNA methylation: an epigenome-wide study of Bangladeshi adults. *Environmental Health Perspectives*, 127(5), 2019. doi: 10.1289/EHP3849.

Gliga, A R; Engström, K; Kippler, M; Skröder, H; Ahmed, S; Vahter, M; Raqib, R, and Broberg, K. Prenatal arsenic exposure is associated with increased plasma IGFBP3 concentrations in 9-year-old children partly via changes in DNA methylation. *Archives of Toxicology*, 92(8):2487–2500, 2018. doi: 10.1007/s00204-018-2239-3.

Green, B B; Karagas, M R; Punshon, T; Jackson, B P; Robbins, D J; Houseman, E A, and Marsit, C J. Epigenome-wide assessment of DNA methylation in the placenta and arsenic exposure in the New Hampshire Birth Cohort Study. *Environmental Health Perspectives*, 124(8):1253–60, 2016. doi: 10.1289/ehp.1510437.

Guo, X; Chen, X; Wang, J; Liu, Z; Gaile, D; Wu, H; Yu, G; Mao, G; Yang, Z; Di, Z; Guo, X; Cao, L; Chang, P; Kang, B; Chen, J; Gao, W, and Ren, X. Multi-generational impacts of arsenic exposure on genome-wide DNA methylation and the implications for arsenic-induced skin lesions. *Environment International*, 119:250–263, 2018. doi: 10.1016/J.ENVINT.2018.06.024.

IARC Working Group. Arsenic, metals, fibres, and dusts. *IARC Monographs on the Evaluation of Chronic Risks to Humans*, 100(Pt C):11–465, 2009.

Kaushal, A; Zhang, H; Karmaus, W J J; Everson, T M; Marsit, C J; Karagas, M R; Tsai, S-F; Wen, H-J, and Wang, S-L. Genome-wide DNA methylation at birth in relation to *in utero* arsenic exposure and the associated health in later life. *Environmental Health*, 16(1):50, 2017. doi: 10.1186/s12940-017-0262-0.

Kile, M L; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Cardenas, A; Wright, R O, and Christiani, D C. Effect of prenatal arsenic exposure on DNA methylation and leukocyte subpopulations in cord blood. *Epigenetics*, 9(5):774–82, 2014. doi: 10.4161/epi.28153.

Milton, A H; Hussain, S; Akter, S; Rahman, M; Mouly, T A, and Mitchell, K. A review of the effects of chronic arsenic exposure on adverse pregnancy outcomes. *International Journal of Environmental Research and Public Health*, 14(6), 2017. doi: 10.3390/ijerph14060556.

Moon, K A; Oberoi, S; Barchowsky, A; Chen, Y; Guallar, E; Nachman, K E; Rahman, M; Sohel, N; D'Ippoliti, D; Wade, T J; James, K A; Farzan, S F; Karagas, M R; Ahsan, H, and Navas-Acien, A. A dose-response meta-analysis of chronic arsenic exposure and incident cardiovascular disease. *International Journal of Epidemiology*, 46(6):1924–1939, 2017. doi: 10.1093/ije/dyx202.

National Research Council. Critical aspects of EPA's and IRIS assessment of inorganic arsenic, interim report. Technical report, National Academies, Washington, DC, 2013.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Sanchez, T; van Geen, A; Mey, J L; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T; Balac, O; Ilievski, V; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folic acid and creatine as therapeutic approaches to lower blood arsenic: a randomized controlled trial. *Environmental Health Perspectives*, 123(12):1294–1301, 2015. doi: 10.1289/ehp.1409396.

Rojas, D; Rager, J E; Smeester, L; Bailey, K A; Drobná, Z; Rubio-Andrade, M; Stýblo, M; García-Vargas, G, and Fry, R C. Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. *Toxicological Sciences*, 143(1):97–106, 2015. doi: 10.1093/toxsci/kfu210.

Röllin, H B; Channa, K; Olutola, B G, and Odland, J Ø. Evaluation of *in utero* exposure to arsenic in South Africa. *Science of the Total Environment*, 575:338–346, 2017. doi: 10.1016/j.scitotenv.2016.10.044.

Smith, A H; Marshall, G; Liaw, J; Yuan, Y; Ferreccio, C, and Steinmaus, C. Mortality in young adults following *in utero* and childhood exposure to arsenic in drinking water. *Environmental Health Perspectives*, 120(11):1527–31, 2012. doi: 10.1289/ehp.1104867.

Steinmaus, C; Ferreccio, C; Acevedo, J; Yuan, Y; Liaw, J; Duran, V; Cuevas, S; Garcia, J; Meza, R; Valdes, R; Valdes, G; Benitez, H; VanderLinde, V; Villagra, V; Cantor, K P; Moore, L E; Perez, S G; Steinmaus, S, and Smith, A H. Increased lung and bladder cancer incidence in adults after *in utero* and early-life arsenic exposure. *Cancer Epidemiology Biomarkers and Prevention*, 23(8):1529–38, 2014. doi: 10.1158/1055-9965.EPI-14-0059.

Steinmaus, C M; Ferreccio, C; Romo, J A; Yuan, Y; Cortes, S; Marshall, G; Moore, L E; Balmes, J R; Liaw, J; Golden, T, and Smith, A H. Drinking water arsenic in northern Chile: high cancer risks 40 years after exposure cessation. *Cancer Epidemiology Biomarkers and Prevention*, 22(4):623–30, 2013. doi: 10.1158/1055-9965.EPI-12-1190.

Tice, R R; Yager, J W; Andrews, P, and Crecelius, E. Effect of hepatic methyl donor status on urinary excretion and DNA damage in B6C3F1 mice treated with sodium arsenite. *Mutation Research*, 386(3):315–34, 1997. doi: 10.1016/S1383-5742(97)00004-5.

U.S. Environmental Protection Agency. Government Performance and Results Act (GPRA) tool-submission year quarter 2018Q3, violation report, 2018. URL `https://obipublic.epa.gov/analytics/saw.dll?PortalPages&PortalPath=/shared/SFDW/_portal/Public&Page=Violation`.

US EPA Office of Water. Drinking water standard for arsenic. Technical report, US Environmental Protection Agency, 2001. URL https://nepis.epa.gov/Exe/ZyPdf.cgi?Dockey=20001XXC.txt.

Vahter, M and Marafante, E. Effects of low dietary intake of methionine, choline or proteins on the biotransformation of arsenite in the rabbit. *Toxicology Letters*, 37(1): 41–46, 1987. doi: 10.1016/0378-4274(87)90165-2.

World Health Organization. Arsenic, 2012. URL http://www.who.int/mediacentre/factsheets/fs372/en/.

Xu, L; Yokoyama, K; Tian, Y; Piao, F-Y; Kitamura, F; Kida, H, and Wang, P. Decrease in birth weight and gestational age by arsenic among the newborn in Shanghai, China. *[Nihon koshu eisei zasshi] Japanese Journal of Public Health*, 58(2):89–95, 2011.

Zhong, Q; Cui, Y; Wu, H; Niu, Q; Lu, X; Wang, L, and Huang, F. Association of maternal arsenic exposure with birth size: a systematic review and meta-analysis. *Environmental Toxicology and Pharmacology*, 69:129–136, 2019. doi: 10.1016/j.etap.2019.04.007.

Chapter 2

Background

2.1 Introduction

Arsenic (As) is naturally-occurring metalloid. Although As comprises an estimated 0.05% of the earth's crust by weight (Fleischer, 1953), As is found ubiquitously in the environment (National Research Council, 1977) and elevated levels of As are present in groundwater in many regions of the world. Known as the "king of poisons" and "poison of kings," the toxicity of As has been known for millennia (Frith, 2013); however, today As exposure continues to affect the health of millions of people worldwide.

The metabolism of ingested inorganic As (InAs) facilitates As elimination in urine. InAs may undergo a series of reduction and oxidative methylation reactions forming monomethyl (MMAs)- and dimethyl (DMAs)-arsenical species. $DMAs^V$ is the least toxic and most readily excreted As species (Petrick et al., 2001). Chronic exposure to As through drinking water and food has been associated with a broad range of health outcomes. Increased risk of adverse health outcomes persists after remediation of As exposure, suggesting that epigenetic dysregulation is a mechanistic link between exposure and health effects.

This chapter will begin with an overview of As exposure, particularly in Bangladesh and the United States (US) (Section 2.2), and its associated health outcomes (Section 2.3 and Section 2.4). This chapter will also describe As metabolism (Section 2.5), one-carbon metabolism, a biological pathway influencing As methylation, and the relationship between one-carbon metabolism nutrients and arsenic methylation capacity (Section 2.6). Finally, potential mechanisms of action of As will be discussed, focusing on epigenetic dysregulation (Section 2.7).

2.2 Sources and extent of arsenic exposure

Arsenic may be released into groundwater through the reductive dissolution of iron and aluminum oxides and the activity of metal-reducing bacteria (Shankar et al., 2014). Anthropogenic sources, including mining, coal combustion, and As-based pesticides, also contribute to elevated concentrations of As in groundwater (Nriagu et al., 2007), although to a lesser extent (Shankar et al., 2014). At least half of the world's population relies on groundwater for drinking (World Water Assessment Programme, 2015). Although exposure to As through drinking water has decreased over the past decades, over 140 million people in more than 70 countries remain exposed to As concentrations greater than 10 μg/L, the World Health Organization (WHO) guideline and US Environmental Protection Agency maximum contaminant level (US EPA Office of Water, 2001; World Health Organization, 2011) (World Health Organization, 2011). The ten countries most severely affected by As exposure and bearing a high proportion of the global burden of As-related morbidity and mortality are Bangladesh, Cambodia, China, India, Myanmar, Nepal, Pakistan, Taiwan, Argentina, and Vietnam (Ravenscroft et al., 2009).

Arsenic-containing foods and beverages can contribute a large proportion of ingested inorganic As (InAs) in areas with low drinking water As concentrations. Rice and rice products are a common source of As exposure. In 2011, rice provided almost 20% of the per capita caloric intake globally (International Rice Research Institute, 2011). Silicon (Si) is necessary for the growth of rice. However, due to the chemical similarity of Si and reduced InAs (As^{III}), which is present in anaerobic conditions such as rice paddies, transporters in rice can readily uptake InAs (Chen et al., 2017). Therefore, rice and rice products may contain elevated levels of As when cultivated with As-contaminated water or on fields treated with As-based pesticides (Potera, 2007). Arsenic concentrations in rice are also affected by cooking rice with As-contaminated

water (Kumarathilaka et al., 2019). Brown rice and brown rice products contain higher inorganic As concentrations due to accumulation in the aleurone layer of the grain (Meharg et al., 2008). The frequency of rice consumption, and consequently As exposure through food, differs among ethnic groups (Food and Drug Administration, 2016). Arsenic exposure through food is a particular concern for infants and children due to rice-based infant products, high per-body-mass food intake, and a more limited diet than adults (Food and Drug Administration, 2016). Additional sources of As exposure include apple juice, red wine, and chicken (Food and Drug Administration, 2017; Nachman et al., 2013).

Arsenic exposure in Bangladesh

Beginning in the 1970s, the United Nations Children's Fund (UNICEF) advocated for the use of tube wells due to high rates of acute gastrointestinal disease among infants and children caused by bacterial contamination of surface water (Smith et al., 2000). By 1997, UNICEF declared that 80% of Bangladesh's population had access to drinking water provided by well and taps, which was thought to be safe. However, also during the 1990s, water testing indicated widespread As contamination of tube wells and health surveys identified As-induced skin lesions affecting over half of patients examined in some areas (Smith et al., 2000). The size of the population affected by As-contaminated drinking water has been estimated by surveys of tube wells. For example, a survey of tube wells conducted between 1995-1999 estimated that 79.9 million people in Bangladesh may be affected by As concentrations exceeding the country's limit of 50 μg/L (Chowdhury et al., 2000). Exposure to drinking water with As concentrations > 10 μg/L contributed to an estimated 43,000 deaths per year nationally in 2009 (Flanagan et al., 2012). Arsenic exposure has been reduced, but in 2013, it was estimated that approximately 40 million people remained exposed to

drinking water As concentrations exceeding 10 μg/L (Bangladesh Bureau of Statistics and United Nation Children's Fund, 2015).

Arsenic exposure in the US

Groundwater As concentrations are elevated in areas of the US including the West, Midwest, and Northeast (Ayotte et al.). In 2001, the US Environmental Protection Agency (EPA) decreased the maximum contaminant level (MCL) from 50 μg/L to 10 μg/L, and required public water systems to become compliant by January 2006 (US Environmental Protection Agency, 2016). Since 2001, the number of people served by public water systems with As concentrations > 10 μg/L has decreased from 13 million to 296,000 (U.S. Environmental Protection Agency 2018). However, the US EPA MCL does not apply to private wells, and an estimated 2.1 million individuals rely on well water with elevated As levels (Ayotte et al., 2017).

Some communities in the US are disproportionately affected by As exposure through drinking water due to geographical location and water source. Arsenic exposure has been measured in the Strong Heart Study (SHS), a population-based prospective cohort of American Indians in Arizona, Oklahoma, and North and South Dakota. Residents of these communities commonly rely of private water well water or small public water systems. Water sampling found elevated As concentrations in public water systems prior to the establishment of the 10 μg/L MCL, particularly in Arizona where As concentrations ranged from < 10 to 61 μg/L (Navas-Acien et al., 2009). Urinary As concentrations were also elevated (median in Arizona = 12.5 μg/g creatinine; Dakotas = 9.1 μg/g creatinine; Oklahoma = 4.4 μg/g creatinine). This study population provides a base on which to test Hypothesis 2 described in **Chapter 5**.

2.3 Health effects of arsenic exposure

Arsenic is a human toxicant and group 1 carcinogen (World Health Organization, 2011). Chronic As exposure has been associated with increased risk of numerous health conditions including skin lesions (melanosis, leukomelanosis, and keratosis), impaired intellectual function, cardiovascular disease, diabetes, inflammation, and cancers including bladder, lung, kidney, liver, skin, and possibly prostate (IARC Working Group, 2009; Moon et al., 2017; National Research Council, 2013) (**Figure 2.1**). Many health effects associated with As exposure have a long latency period and an elevated risk of disease persists even decades after exposure has been reduced or ended. In Northern Chile, where people were exposed to As-contaminated drinking water for a very distinct temporal period, exposure isolated to infancy or early life has been associated with elevated lifetime risk of As-related health outcomes such as lung and bladder cancer (Smith et al., 2012; Steinmaus et al., 2013, 2014).

Arsenic's effects on nervous system development are a particular concern for child health. Reviewed by Tyler and Allan 2014, multiple epidemiological studies have demonstrated adverse effects of childhood As exposure on intelligence, cognitive skills, visual perception, and mental health (Wasserman et al., 2011, 2016). Chronic exposure to medium levels of As appear to have stronger associations with neurodevelopment outcomes than acute high-level exposure during periods of development.

2.4 *In utero* arsenic exposure and birth outcomes

In utero As exposure may have adverse effects of fetal development. Decreased birth size, which may reflect the intrauterine environment and development, has been associated with increased morbidity and mortality in adulthood (Kajantie et al., 2005; Knop et al., 2018). Arsenic crosses the placenta, and maternal and cord blood As

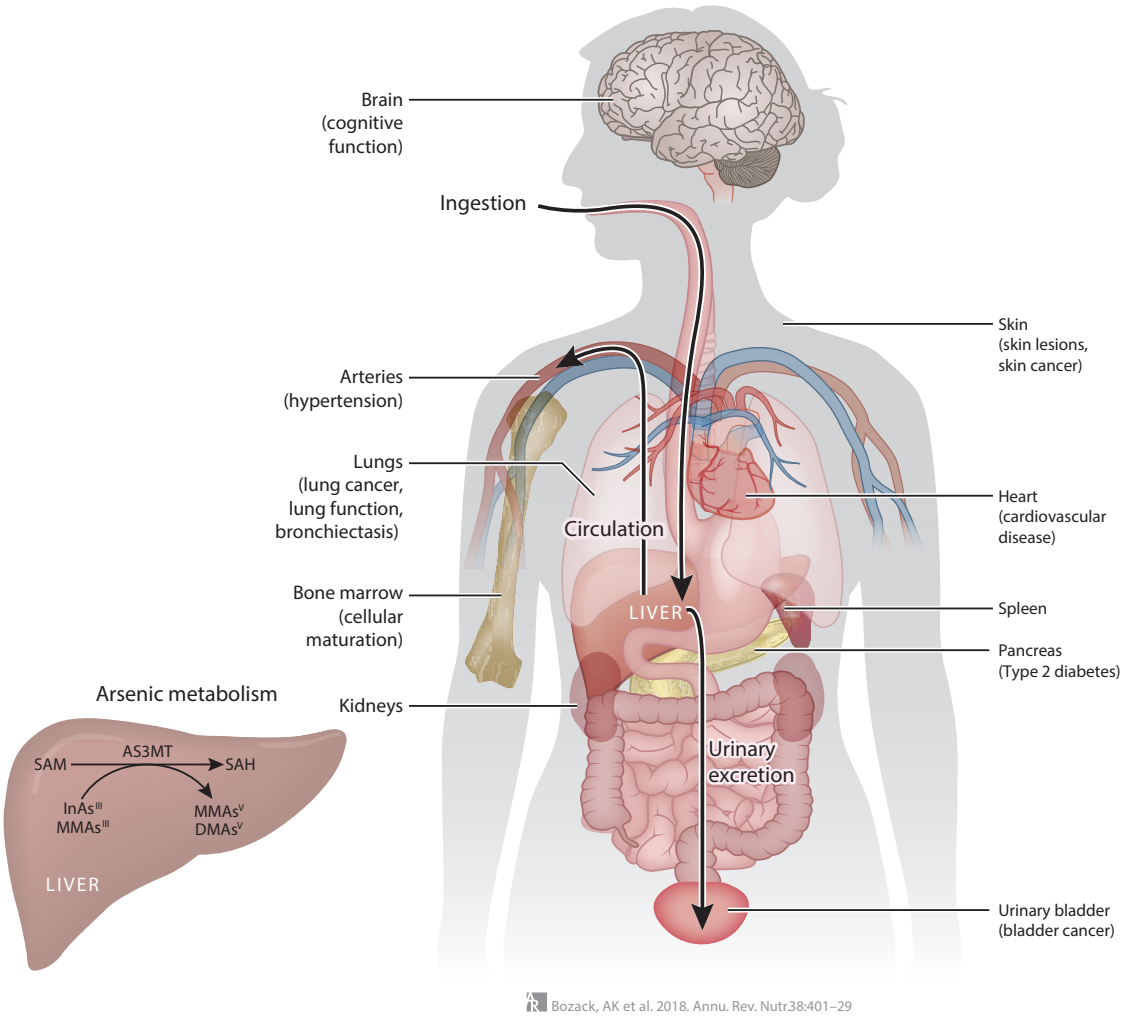


Figure 2.1: Arsenic metabolism, target tissues, and comorbidities. Chronic As exposure has been associated with increased risk of skin lesions (melanosis, leukomelanosis, and keratosis), cardiovascular disease, hypertension, impaired intellectual function, inflammation, diabetes, and cancers. Ingested As accumulates in multiple tissues, including the spleen, liver, lungs, kidneys, bladder, skin, and bone marrow. (Inset) AS3MT is predominantly expressed in the liver, although AS3MT mRNA has also been detected in the kidneys, adrenal gland, bladder, heart, and brain. Abbreviations: AS3MT, arsenic-3-methyltransferase; $DMAs^{V}$, dimethylarsinic acid; $InAs^{III}$, arsenite; MMAsIII, monomethylarsonous acid; $MMAs^{V}$, monomethylarsonic acid; SAH, *S*-adenosylhomocysteine; SAM, S-adenosylmethionine (Bozack et al., 2018c).

concentrations are highly correlated (Concha et al., 1998; Hall et al., 2007). The teratogenic effects of *in utero* As exposure have been demonstrated in rodent studies. These include decreased fetal weight (Ferm and Hanlon, 1985; Hill et al., 2008; Kozul-Horvath et al., 2012; Nagymajtényi et al., 1985), delayed fetal development (Moore et al., 2019), fetal malformations, (Hood, 1972; Hood and Bishop, 1972) and neural tube defects (Hill et al., 2008; Hood, 1972; Hood and Bishop, 1972; Morrissey and Mottet, 1983).

Birth outcomes associated with *in utero* As exposure have also been assessed in epidemiological studies, as reviewed by Bloom et al. (2014); Milton et al. (2017); Vahter (2009), although findings have been inconsistent. Studies of the associations between As exposure and outcomes of birth weight and gestational age are summarized in Appendix Table 2.1 and Appendix Table 2.2, respectively. A recent review by Milton et al. focused on epidemiological studies investigating the association between maternal exposure to As in drinking water and adverse pregnancy outcomes including spontaneous abortion, stillbirth, preterm birth, and neonatal mortality. Although findings have differed among studies, the review concluded that there is "consistent and convincing evidence" of the association between as high As exposure and increased risk of spontaneous abortion and stillbirth, but an insufficient number of studies addressing of the associations between As exposure and neonatal death and preterm birth (Milton et al., 2017).

Milton et al. also summarized studies of the association between *in utero* As exposure and low birth weight. Among 11 studies included in the review, six reported elevated risk of low birth weight with increased As exposure and four reported a null association (Milton et al., 2017). More recently, Zhong et al. conducted a meta-analysis of epidemiological studies of the association between maternal As exposure and anthropometric measurements at birth (Zhong et al., 2019). Twelve studies of the association between As exposure and birth weight were included in the meta-analysis resulting in a

summary regression coefficient of -25.0 g (95% CI: -41.0, -9.0). Similarly, meta-analysis results indicated a negative association between *in utero* As exposure and birth length (β = -0.12 cm; 95% CI: -0.17, -0.07) and head circumference (β = -0.12 cm; 95% CI: -0.24, -0.01).

The effect of As exposure on intrauterine growth may be modified by factors including exposure level, birth size, and infant sex. In a birth cohort in Bangladesh, Rahman et al. found a significant interaction between maternal urinary As concentrations and exposure levels ($<$ 100 μg/L vs. $\leq$ 100 μg/L), resulting in a negative effect of As exposure on birth weight at low levels, but not high, levels of exposure (Rahman et al., 2008). In a separate Bangladeshi cohort, mediation analysis of the association between *in utero* As exposure and birth weight indicated a significant independent association between maternal water As and birth weight only among infants in the 10th to 20th percentiles of birth weight (10th percentile: β = -28.0 g, 95% CI: -43.8, -9.9; 20th percentile: β = -14.9 g, 95% CI = -30.3, -1.7) (Rahman et al., 2017). In addition, in a birth cohort in China, analyses stratified by sex indicated a significant negative association between maternal bAs and birth weight among male infants (β = -354.41, 95% CI = -677.53, -31.28), but not among female, infants (Xu et al., 2011).

Although many studies of *in utero* As exposure have focused on the effects of As on birth weight, low birth weight is caused by short gestation and/or intrauterine grown restriction (Wardlaw et al., 2004). These factors have different effects on health outcomes: shorter gestation is associated with increased risk of infant mortality, morbidity, and disability; restricted intrauterine growth is associated with decreased growth in childhood and increased morbidity in adulthood. Distinguishing between etiological factors is therefore important to understanding the health implications of *in utero* As exposure.

Several studies have evaluated birth weight and gestational age as indepen-

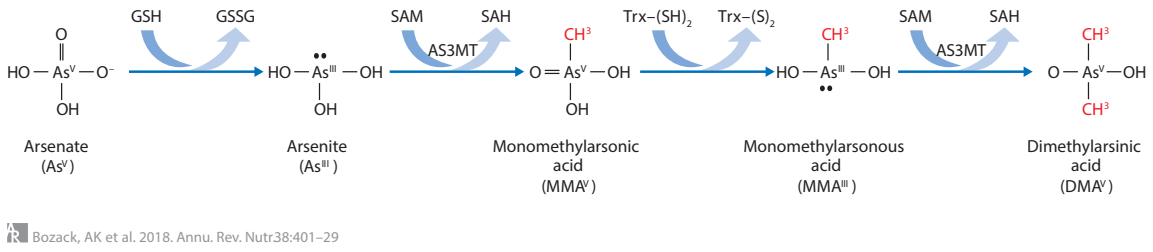
dent outcomes associated with *in utero* As exposure, finding no significant association with either outcome (Bloom et al., 2016; Freire et al., 2019; Sun et al., 2019), or a significant negative association with both outcomes (Xu et al., 2011). To help elucidate the relationship between As exposure, birth weight, and gestational age, in a birth cohort in Bangladesh, David Christiani's group has conducted mediation analyses including both birth outcomes (Kile et al., 2015; Lin et al., 2019; Wei et al., 2017; Rahman et al., 2017). These results suggest that the observed association between *in utero* As exposure and birth weight is fully mediated by a shorten gestational period: in structural equation models (SEMs), the relationship between maternal drinking water As levels and birth weight (β = -19.17, 95% CI: -24.64, -13.69) was mediated through gestational age (β = -17.37, 95% CI: -22.77, -11.98) and maternal weight gain (β = -1.80, 95% CI: -3.72, 0.13) (Kile et al., 2015).

Although epidemiological studies of the relationship between *in utero* As exposure and birth outcomes have not provided consistent results, overall there is a trend toward a negative association between exposure and birth size and gestational length. Conclusions from systematic literature reviews and a meta-analysis strengthen the evidence of the association between *in utero* As exposure and adverse birth outcomes. Differences in results of epidemiological studies may be due in part is biases in study design, differences in level of exposure, methods of exposure assessment, sample size, and statistical analyses performed. Further research is needed to fully understand the association between *in utero* As exposure and birth outcomes. The associations between *in utero* As exposure, birth weight and gestational age, and mediation by epigenetic dysregulation be discussed further in **Chapters 6 and 7**.

2.5 Arsenic methylation

Section 2.5, Arsenic methylation, was published in Bozack AK*, Saxena R*, Gamble M V. 2018. Nutritional Influences on One-Carbon Metabolism: Effects on Arsenic Methylation and Toxicity. *Annu Rev Nutr.* 2018. 38:40129. (Bozack et al., 2018c). (*these authors contributed equally to this article).

Arsenic in drinking water is predominantly inorganic arsenate ($InAs^{V}$) or arsenite ($InAs^{III}$), but once ingested, it undergoes methylation in a process that facilitates urinary As elimination (Buchet et al., 1981). In 1945, Challenger reported a model in which InAs undergoes alternate reduction and oxidative methylation reactions (Challenger, 1945), illustrated in Figures 2.1 and 2.2. Briefly, $InAs^{III}$ is methylated by arsenic-3-methyltransferase (AS3MT) (Lin et al., 2002), using *S*-adenosylmethionine (SAM) as the methyl donor, to form monomethylarsonic acid ($MMAs^{V}$) (Del Razo et al., 2001). $MMAs^{V}$ is then reduced to monomethyl-arsonous acid ($MMAs^{III}$), an intermediate with very high cytotoxicity and genotoxicity (Petrick et al., 2001, 2000; Styblo et al., 2000). $MMAs^{III}$ is subsequently methylated by AS3MT to form dimethylarsinic acid ($DMAs^{V}$) (Challenger, 1945; Lin et al., 2002; Thomas et al., 2004). $DMAs^{V}$ is rapidly excreted in urine and is considerably less toxic than $MMAs^{III}$, $InAs^{III}$, or $InAs^{V}$ (Styblo et al., 2000). While other identified methyltransferase enzymes are capable of methylating As (Zakharyan et al., 1999, 1995), AS3MT catalyzes these methylation reactions with a Km in the nanomolar range, indicating that it is the most physiologically relevant enzyme for As methylation (Lin et al., 2002). Glutathione (GSH) may increase the speed of the reduction steps, influence the activity of AS3MT, and sequester As. Hayakawa et al. 2005 proposed a pathway in which AsGSH complexes are substrates for AS3MT. Using a mathematical model based on the known biochemistry of As derived from cellular and experimental studies, our group found that the Challenger pathway of As methylation, along with the GSH effects, is sufficient to understand and predict ex-



Bozack, AK et al. 2018. Annu. Rev. Nutr38:401–29

Figure 2.2: Arsenic methylation. According to the Challenger pathway (Challenger, 1945), AS3MT catalyzes the oxidative methylation of arsenite using SAM as the methyl donor, forming $MMAs^{V}$ and SAH. $MMAs^{V}$ is then reduced to $MMAs^{III}$ before a subsequent oxidative methylation step, yielding $DMAs^{V}$ and SAH. Abbreviations: AS3MT, arsenic-3-methyltransferase; $DMAs^{V}$, dimethylarsinic acid; $MMAs^{III}$, monomethylarsonous acid; $MMAs^{V}$, monomethylarsonic acid; SAH, *S*-adenosylhomocysteine; SAM, *S*-adenosylmethionine. (Bozack et al., 2018c).

perimental data (Lawley et al., 2014). Additional experimental studies are needed to determine the relative roles that alternative pathways play in As methylation.

AS3MT is predominantly expressed in the liver, but *AS3MT* mRNA has also been detected in the kidneys, adrenal gland, bladder, heart, and brain (Lin et al., 2002). Work by Thomas's group (Drobna et al., 2009; Hughes et al., 2010) using *AS3MT* knockout mice illustrates the profound role of AS3MT in As elimination, as mice deficient in *AS3MT* accumulate a body burden of As that is 16-20 times greater than wild-type mice and exhibit severe systemic toxicity and early death.

There are controversies surrounding the influence of As metabolism on toxicity owing in part to analytical challenges. For example, because trivalent arsenicals are readily oxidized to pentavalent forms by environmental oxygen, it is difficult to distinguish between the valence states of As metabolites, and the potential for artifact is high. Because of this limitation, most human studies report the percentages of total As $\%InAs^{III+V}$, $\%MMAs^{III+V}$, and $\%DMAs^{III+V}$. Some studies have provided conflicting information about the portion of $DMAs^{V}$ versus $DMAs^{III}$ that may be present in urine (Mandal et al., 2001; Del Razo et al., 2001; Valenzuela et al., 2004). This reported discrepancy may be due to the high reactivity of $DMAs^{III}$, which can be rapidly oxidized

to $DMAs^{V}$ (Kobayashi and Hirano, 2008). Additionally, one chromatographic protocol that treats $DMAs^{V}$ with metabisulfite and thiosulfate can inadvertently generate thio-$DMAs^{V}$ (Hansen et al., 2004), a relatively minor arsenical species (5% of total As) that has been identified in human urine (Raml et al., 2007).

Arsenic Methylation and Toxicity

Arsenic metabolites vary considerably in their toxicity. Trivalent arsenicals have greater cytotoxicity and genotoxicity than pentavalent forms. Using real-time cell sensing in two human cell lines, Le's group (Moe et al., 2016), in agreement with the work of others (Petrick et al., 2000), reported that $MMAs^{III}$ is the most cytotoxic As metabolite, followed by $DMAs^{III}$, $InAs^{III}$, $InAs^{V}$, $MMAs^{V}$, $DMAs^{V}$, and the chicken feed additive Roxarsone. Studies utilizing hamster cells from a variety of tissues including kidney and heart demonstrate that $MMAs^{III}$ is the most cytotoxic arsenical species in all cell types tested (Petrick et al., 2001). In epidemiological studies, a higher percentage of $MMAs^{III+V}$ in urine has been associated with increased risk for bladder, breast, lung, and skin cancer, as well as skin lesions, peripheral vascular disease, hypertension, atherosclerosis (Figure 2.3), and decreased birthweight (Ahsan et al., 2007; Chen et al., 2003a,b; Gilbert-Diamond et al., 2013; Hsueh et al., 1997; Huang et al., 2006, 2008a; Laine et al., 2015; Lindberg et al., 2008; López-Carrillo et al., 2014; Melak et al., 2014; Pu et al., 2007; Steinmaus et al., 2006, 2010; Tseng et al., 2005; Wu et al., 2006; Yu et al., 2000). In contrast, the risk of metabolic syndrome and diabetes has been negatively associated with %MMAs in urine (Chen et al., 2012; Kuo et al., 2017). Most, but not all, of these studies used prevalent cases and therefore cannot establish temporality or rule out the possibility of reverse causation, i.e., that having a disease influences the ability to methylate As. However, in a recent nested case-control study of As-induced skin lesions, we found that participants falling into the lowest tercile of %DMAs in

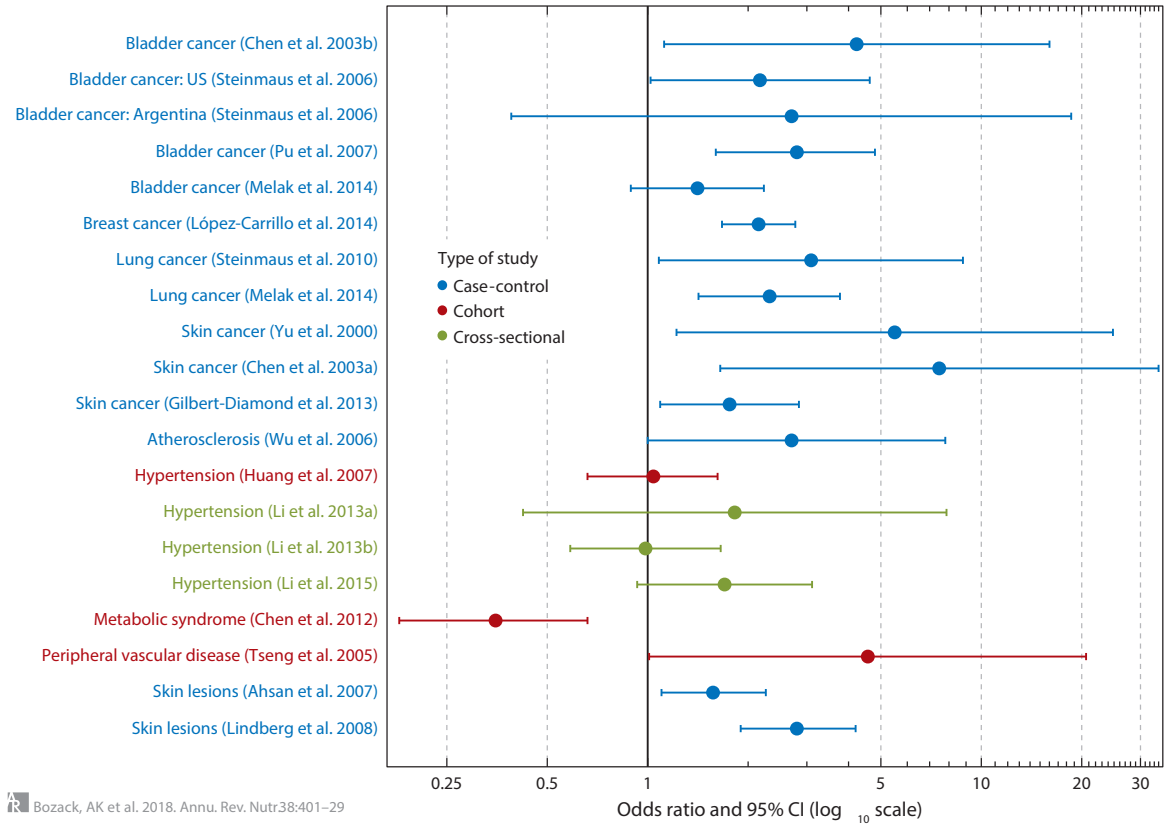


Figure 2.3: Summary plot of odds ratios and 95% CIs for health outcomes reported to be associated with %MMAs in urine (Ahsan et al., 2007; Chen et al., 2012, 2003a,b; Gilbert-Diamond et al., 2013; Hsueh et al., 1997; Huang et al., 2006, 2008a; Laine et al., 2015; Lindberg et al., 2008; Li et al., 2013b,a, 2015; López-Carrillo et al., 2014; Melak et al., 2014; Pu et al., 2007; Steinmaus et al., 2006, 2010; Tseng et al., 2005; Wu et al., 2006; Yu et al., 2000). Abbreviations: %MMAs: percent monomethyl-arsenical species; CI: confidence interval. (Bozack et al., 2018c).

urine were at higher risk for development of skin lesions 2-7 years later (Niedzwiecki et al., 2018).

Arsenic methylation capacity differs between species, individuals, and populations. In Bangladesh, members of our group have found that genetic variation in the *AS3MT* gene is a strong genetic predictor of As methylation capacity (Gao et al., 2015; Jansen et al., 2016). Single-nucleotide polymorphisms (SNPs) in *AS3MT* have also been associated with the proportion of As metabolites in urine in other populations

(Agusa et al., 2009; Balakrishnan et al., 2017; Engström et al., 2011; Hernández et al., 2008a,b), as well as As-related health outcomes such as skin lesions and skin cancer (Engström et al., 2015; Gao et al., 2015). In addition, epigenetic regulation may influence As methylation capacity. In a study of Argentinean women, the *AS3MT* haplotype was found to be associated with the methylation status of the *AS3MT* gene – which influenced AS3MT gene expression – and with As methylation capacity (Engström et al., 2013).

2.6 One-carbon metabolism and nutritional influences on arsenic methylation

(**Section 2.6, One-carbon metabolism and nutritional influences on arsenic methylation**, was published in Bozack AK*, Saxena R*, Gamble M V. 2018. Nutritional Influences on One-Carbon Metabolism: Effects on Arsenic Methylation and Toxicity. *Annu Rev Nutr.* 2018. 38:40129. (Bozack et al., 2018c). (*these authors contributed equally to this article).

Directly relevant to **Chapters 3 and 4** on the influence of one-carbon metabolism (OCM) nutrients on As methylation, this section will describe the OCM pathway. It will also provide an overview of current evidence of the association between OCM-related nutrients and As methylation capacity.

One-carbon metabolism

Methylation reactions are dependent on the methyl donor SAM, a critical cosubstrate in OCM (Figure 2.4). Several micronutrients including folate, vitamin B_{12} (cobalamin), betaine, choline, riboflavin, and vitamin B_6 (pyridoxal phosphate) play critical roles in OCM. FA is a synthetic form of folate used in food fortification. FA

must be reduced to 5-methyltetrahydrofolate (5-methyl-THF) by dihydrofolate reductase to become metabolically active in OCM. In the folate cycle, a one-carbon unit is transferred from serine to tetrahydrofolate (THF) to form essential role in pulling the pathway forward. Despite the fact that multiple micronutrients influence OCM, folate nutritional status appears to be the primary determinant of circulating homocysteine concentrations in most studies of adults. For example, in a cross-sectional survey, we observed a high prevalence of high homocysteine in Bangladesh, and plasma folate and B_{12} were found to explain 15% and 5% of the variance in homocysteine, respectively (Gamble et al., 2005a).

OCM and Arsenic Metabolism in Animal Models

Early studies in animal models provided experimental evidence that nutritional regulation of OCM influences the methylation and toxicity of As. In 1987, Vahter and Marafante (Vahter and Marafante, 1987) reported that rabbits fed diets deficient in methyl donors (methionine, choline, or protein) had significantly lower urinary excretion of total As and DMAs, and increased As retention in tissues; similar results were reported by others in mice (Tice et al., 1997). Finnell's group conducted an elegant series of studies on neural tube defects (NTDs) that employed mice nullizygous for several folate-binding proteins involved in cellular uptake of folate from the circulation (e.g., Folbp-1 and -2) and/or enterocytes (reduced folate carrier). Mice were injected with sodium arsenate early in gestation during critical periods for neural tube closure. For all genotypes studied, dietary folate deficiency caused a reduction in urinary excretion of DMAs, and Folbp-2 null mice were more susceptible to As-induced NTDs (Spiegelstein et al., 2003, 2005; Wlodarczyk et al., 2001). While As exposure has not yet been solidly linked to NTD risk in humans, studies are under way, and a recent study in Bangladesh reported that As exposure reduced the efficacy of FA in preventing

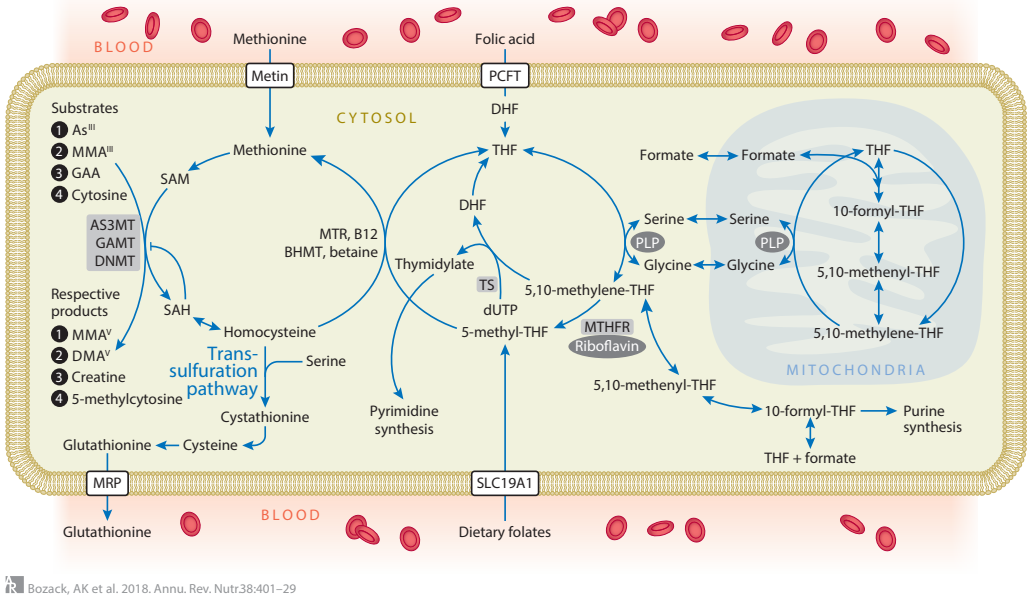


Figure 2.4: One-carbon metabolism. FA, arising from fortified foods or nutritional supplements, is reduced to DHF and THF by dihydrofolate reductase. Serine hydroxymethyl-transferase transfers one-carbon units from serine to THF, with PLP as a coenzyme, forming 5,10-methylene-THF. This is either used for the synthesis of thymidylate or reduced to 5-methyl-THF. Dietary folates can enter one-carbon metabolism as 5-methyl-THF. The methyl group of 5-methyl-THF is transferred to homocysteine in a reaction catalyzed by MTR and utilizing B_{12} as a cofactor, generating methionine and THF. Alternatively, in the liver, betaine can donate a methyl group for the remethylation of homocysteine in a reaction catalyzed by BHMT. Methionine adenosyltransferase enzymes activate methionine to form SAM the methyl donor for numerous acceptors, including arsenicals, GAA (the precursor to creatine), and DNA in reactions that involve substrate-specific methyltransferase enzymes. These methylation reactions generate the methylated products and SAH, a potent product inhibitor of most methyltransferases. SAH is hydrolyzed to generate homocysteine, which is either remethylated to regenerate methionine or directed to the transsulfuration pathway and ultimately glutathionine synthesis. Abbreviations: AS3MT, arsenic-3-methyltransferase; BHMT, betaine homocysteine methyltransferase; DHF, dihydrofolate; $DMAs^V$, dimethylarsinic acid; DNMT, DNA methyltransferases; FA, folic acid; GAA, guanidinoacetate; GAMT, guanidinoacetate-Nmethyltransferase; $InAs^{III}$, arsenite; $MMAs^{III}$, monomethylarsonous acid; $MMAs^V$, monomethylarsonic acid; MRP, multidrug resistance proteins; MTR, methionine synthase; PCFT, proton coupled folate transporter; PLP, pyridoxal phosphate; SAH, textitS-adenosylhomocysteine; SAM, *S*-adenosylmethionine; SLC19A1, solute carrier family 19 (folate transporter), member 1, also known as RFC1; THF, tetrahydrofolate; TS, thymidylate synthetase. (Bozack et al., 2018c).

NTDs (Mazumdar et al., 2015).

The utility of rodent models in understanding the effects of OCM-related micronutrients on As methylation capacity is limited by the facts that (a) there are profound differences between species in As metabolism efficiency (both mice and rats are extremely efficient in As methylation); (b) animals are less prone to developing As-related cancers than are humans; and (c) rodents are much less prone to developing folate deficiency than are humans due to coprophagia. In addition, it is challenging to mimic chronic, often decades-long, lower-dose population-based As exposure levels using rodent models.

Human Studies on Folate and Arsenic Metabolism and Toxicity

Early human data suggesting a role for folate in As toxicity came in the form of isolated case reports. For example, an interesting case study of a girl deficient in MTHFR, the enzyme responsible for the reduction of 5,10-methylene-THF to 5-methyl-THF, developed severe symptoms of As toxicity following exposure to an As-containing pesticide, while no other exposed family members were affected (Brouwer et al., 1992). Several studies on nutrition and As are summarized in Appendix Table 2.3. For example, in 2002, Smith's group Chung et al. (2002) reported a cross-sectional study of 11 families exposed to As-contaminated drinking water in Chile (N = 44) in which the correlations of As methylation patterns between fathers and mothers were low, but they increased substantially with adjustment for folate and homocysteine, indicating that folate and homocysteine are sources of variation in As methylation. Two years later, Smith's group (Mitra et al., 2004) reported a nested case-control study of As-related skin lesions (keratosis and melanosis) in a population with very high water As concentrations in West Bengal, India. While they did not measure As metabolites, there was a positive

trend in risk for skin lesions associated with lower quintiles of folate intake (P for trend = 0.006). In a subsequent case-control study of this population, diet was assessed using a 24-h recall, and the concentrations of selected micronutrients were measured in serum; %MMAs was observed to be significantly lower among participants in the highest tercile of serum folate compared to the lowest tercile (Basu et al., 2011). This same group conducted another study employing dietary questionnaires in As-exposed regions of the western United States, in which they found that participants falling in the lower quartiles for dietary protein, iron, zinc, and niacin had higher %MMAs and lower %DMAs in urine compared to the highest quartile (Steinmaus et al., 2005). While they found no significant associations with dietary folate intake, the study was conducted years after mandatory fortification of cereals, breads, pastas, flours, and other grain products in the US food supply with FA; therefore, all of the study participants likely had fairly high folate intake.

Our group has conducted a series of studies in Bangladesh, a population with chronic As exposure, to better characterize the interconnections between folate, As metabolism, and As toxicity.

We first conducted a cross-sectional study of As-exposed adults in Bangladesh, in which we observed that plasma folate was negatively associated with %MMAs and positively associated with %DMAs Gamble et al., 2005b. We then studied the effect of 400 μg FA per day, the US recommended dietary allowance, on As metabolism in a randomized, double-blind, placebo-controlled trial among 200 folate-deficient (plasma folate $<$ 9 nmol/L) Bangladeshi adults. After 12 weeks of supplementation, the treatment group had a significantly larger increase in %DMAs and decreases in %InAs and %MMAs in urine relative to the placebo group; treatment effects were observed as early as one week post-intervention (Gamble et al., 2006). In addition, FA supplementation lowered blood As (bAs) by 14% and blood MMAs by 22% (Gamble et al., 2007). More

recently, in the Folic Acid and Creatine Trial (FACT) conducted among 622 As-exposed Bangladeshi adults selected independent of folate status, a larger decrease in bAs was observed in the treatment group receiving 800 μg FA per day for 12 weeks relative to the placebo group (Peters et al., 2015a). In addition, increases in As methylation capacity were observed among the treatment groups receiving 800 μg FA per day and 400 μg FA per day relative to placebo as measured by the change in %InAs, %MMAs, and %DMAs in urine between baseline and week 12 (Bozack et al., 2018b) (Reported in detail in **Chapter 3**). FACT included a 12-week wash-out period during which half of the participants in the FA treatment groups were switched to placebo to examine the effects of cessation of supplementation on As methylation. Arsenic metabolites reverted to pre-intervention levels 12 weeks after FA supplementation was discontinued, highlighting the importance of maintaining adequate folate nutritional status over time. This has important policy implications, as prolonged maintenance of FA effects may be more readily achieved through food FA-fortification programs than through recommendations for over-the-counter FA supplements, as the latter has limited long-term compliance.

In 2009, we reported the results of a nested case-control study of skin lesions (274 cases and 274 controls) in Bangladesh in which we found that low folate, hyperhomocysteinemia, and low urinary creatinine were associated with risk for subsequent development of skin lesions after controlling for age, urinary As, and use of betel nut (a mild stimulant and known carcinogen commonly chewed by Bangladeshi adults) (Pilsner et al., 2009). These findings are consistent with those of an analysis of dietary folate and skin lesions in India by Smith and colleagues (Melak et al., 2014). Similar results were found in relation to urothelial carcinoma. In a casecontrol study of 177 cases and 488 controls in a population in Taiwan exposed to low concentrations of As in drinking water, higher %DMAs in urine and higher plasma folate concentrations

were associated with a decreased risk for urothelial carcinoma. Furthermore, a significant interaction was observed between urinary As methylation profiles and plasma folate in affecting urothelial carcinoma risk (Huang et al., 2008b). More recently, in a larger nested case-control study (N = 876 cases and 876 controls) of gene × nutrition × environment interactions, hyperhomocysteinemia and lower %DMAs in urine were both associated with increased risk for development of skin lesions 27 years later (Niedzwiecki et al., 2018). We also found *TYMS* rs1001761 was associated with increased skin lesion risk at water As exposure > 50 μg/L. The latter finding highlights a potential role of OCM in As toxicity independent of As methylation. The *TYMS* gene encodes thymidylate synthetase, which utilizes 5,10-methylene-THF for the methylation of 2-deoxy-uridine-5-monophosphate (dUMP) to 2-deoxy-thymidine-5-monophosphate (dTMP) and is critical for DNA synthesis and repair (Figure 2.4) (Carreras and Santi, 1995). These findings raise the possibility that DNA damage involving thymidylate synthetase may be a mechanism of As toxicity at higher As concentrations. Consistent with this hypothesis, a recent study from the Stover group identified thymidylate biosynthesis as a sensitive target for As at levels observed in human populations (As trioxide at 0.5 μM, equivalent to 75 μg/L As in water) (Kamynina et al., 2017). The study found that As exposure to cell cultures impaired folate-dependent dTMP biosynthesis, resulting in uracil misincorporation into DNA and genomic instability. Further, folate deficiency exacerbated the impact of As on uracil misincorporation and genomic instability, providing a potential additional mechanism linking folate deficiency to skin lesion risk (Kamynina et al., 2017).

Recent, large epidemiological studies have confirmed our findings relating folate to As methylation, including in populations with lower As exposure. In a cross-sectional study of 1,027 women in Mexico with a median urinary As of 25.9 μg/g creatinine, micronutrient intake was estimated using a food frequency questionnaire;

folate intake was associated with significantly lower %InAs and higher methylation ratio of DMAs/InAs (López-Carrillo et al., 2016). In addition, in an analysis of the 2003-2004 National Health and Nutrition Examination Survey, dietary folate intake was negatively associated with urinary %InAs and positively associated with %DMAs in unadjusted models, and red blood cell folate was negatively associated with %InAs in adjusted models (Kurzius-Spencer et al., 2017). In the SHS, a cohort study of American Indian adults with low to moderate As exposure, high combined intake of folate and B_6, as estimated by a food frequency questionnaire, was negatively associated with %InAs in urine and positively associated with %DMAs (Spratlen et al., 2017). The significant associations between dietary intake of folate and As metabolism in these studies differ from the results of a dietary intake analysis by members of our group that reported no association with dietary folate intake in Bangladesh (Heck et al., 2007). However, the prolonged cooking times traditionally used in Bangladesh can degrade folate in foods, making it more difficult to accurately measure true folate intake. Such considerations highlight the importance of using circulating folate concentrations, particularly in this region of the world.

Several studies relating OCM to As methylation have been conducted in pregnant women. During pregnancy, OCM influences exposure of the fetus to As, as InAs, MMAs, and DMAs are all transported through the placenta. Furthermore, As concentrations and As metabolites are similar in maternal and umbilical cord blood (Concha et al., 1998; Hall et al., 2007). OCM is altered during pregnancy owing to the demands of fetal development. For example, maternal plasma folate levels change dramatically over the course of pregnancy. Additionally, endogenous choline synthesis is induced by estrogen and upregulated during pregnancy and lactation (Zeisel, 2009) (Zeisel 2009).

The relationship between OCM-related micronutrients and As metabolism in

pregnant women has been investigated in several studies. In a cross-sectional study of women at 14 weeks gestation in Bangladesh (N = 753), Vahters group reported an inverse association between plasma folate terciles and urinary %InAs. Women who were deficient in folate, vitamin B_{12}, and zinc had significantly higher %InAs and lower primary methylation index (MMAs/InAs) compared to women who were not deficient in any of the three nutrients (Li et al., 2007). Vahter's group conducted a subsequent longitudinal study of Bangladeshi women (N = 324) assessed at gestational weeks 8, 14, and 30. They observed significant negative associations between gestational week and urinary %InAs and %MMAs, and a significant positive association between gestational age and %DMAs; however, neither plasma folate nor B_{12} were associated with the proportions of As metabolites (Gardner et al., 2011). The conclusion that these micronutrients have little effect on As methylation during pregnancy may be complicated by changes in OCM during pregnancy. In addition, all women received a daily supplement of 400 μg FA beginning at week 14, which may have impacted As metabolism.

Impact of Other OCM-Related Nutrients on Arsenic Methylation

Although the association between folate and As methylation has been broadly studied, other micronutrients involved in OCM (Figure 2.4) may influence As methylation capacity. Below, we summarize research on the associations between creatine, vitamin B_{12}, choline, and betaine and As methylation profiles in humans. These OCM micronutrients are less widely examined than folate in terms of their relationship to As methylation, and are potential avenues for future research directions.

Creatine. Creatine is a nitrogenous organic acid that is present in foods and is also synthesized endogenously. Our group (Bozack et al., 2018b; Gamble et al.,

2005b, 2006; Hall et al., 2009) and others (Basu et al., 2011; Kile et al., 2009) have consistently reported that urinary creatinine, a degradation product of creatine, is a strong predictor of As methylation capacity; it is positively associated with %DMAs in urine and negatively associated with %InAs. The synthesis of creatine, the precursor of creatinine, consumes approximately 50% of all SAM-derived methyl groups (Mudd and Poole, 1975; Stead et al., 2006). In omnivores, roughly half of creatine requirements are met through dietary intake of creatine, primarily from meat (Brosnan et al., 2011). Urinary creatinine concentrations therefore reflect both dietary creatine intake and endogenous creatine synthesis. Creatinine is also commonly used in urinalyses to adjust for hydration status. Increases in circulating creatine concentrations, e.g., from dietary intake, lower creatine biosynthesis by inhibiting synthesis of guanidinoacetate (GAA), the precursor of creatine; in rodents, this has been shown to spare methyl groups and lower homocysteine (Deminice et al., 2008; Guthmiller et al., 1994; Stead et al., 2001; Taes et al., 2003). We hypothesized that creatine supplementation may also spare methyl groups and thereby facilitate the methylation of As, and may underlie the observed associations between the proportion of urinary As metabolites and urinary creatinine.

We tested this hypothesis in our FACT study. Creatine supplementation for 12 weeks at 3 g per day (roughly 1.5 times the normal daily creatine turnover for a 70-kg male) lowered GAA as predicted, illustrating that creatine supplementation inhibited GAA synthesis (Peters et al., 2015b). Also, the mean decrease in urinary %MMAs in the creatine treatment group exceeded that of the placebo group at weeks 6 and 12 ($P < 0.05$); however, creatine supplementation did not affect the change in %InAs or %DMAs (Bozack et al., 2018b). Creatine treatment effects may have been tempered by long-range allosteric regulation of OCM. Further research is needed to understand the strong cross-sectional associations between urinary creatinine and As methylation in

previous studies. One possibility is that urinary creatinine is somehow related to renal tubular reabsorption of InAs, a topic that is understudied in the scientific literature. Further results of this trial are reported in **Chapters 3 and 4**.

Vitamin B_{12} (cobalamin). Vitamin B_{12} has been inconsistently linked to As metabolism in human studies, but has not been studied in animal models. In an early cross-sectional study in Bangladesh in which we oversampled B_{12}-deficient individuals, we found plasma B_{12} concentrations to be inversely associated with %InAs in urine and positively associated with %MMAs, with no association with %DMAs (Hall et al., 2009). We found it difficult to reconcile a mechanism whereby B_{12} would facilitate mono- but not dimethylation of As. To test whether this might have been a spurious finding, we subsequently analyzed the relationship between B_{12} and As metabolites in blood and urine in FACT and the Folate and Oxidative Stress Study (FOX). In baseline samples from the FACT study, there were no significant associations between B_{12} and As metabolites in urine (nonsignificant correlations were negative for %InAs and %MMAs and positive for %DMAs), and all associations with bAs metabolite concentrations were negative. In FOX, while B_{12} was positively associated with %MMAs in urine ($r_{Spearman}$ = 0.12, P = 0.02) the correlation with MMAs concentrations in blood was null ($r_{Spearman}$ = -0.02, P = 0.76) (M.V. Gamble, unpublished data). In an analysis of the 2003-2004 National Health and Nutrition Examination Survey, in adjusted models, dietary B_{12} intake was positively associated with urinary %InAs and negatively associated with %DMAs (Kurzius-Spencer et al., 2017), but other studies have reported null results (Spratlen et al., 2017) or contrasting directions of association (López-Carrillo et al., 2016). To date, no studies investigating the treatment effects of B_{12} supplementation on As methylation have been published. Additional studies are needed to better understand the relationship between B_{12} and As methylation.

Choline and betaine. The remethylation of homocysteine to methionine, crit-

ical for the synthesis of SAM, can be catalyzed either by methionine synthase, which receives a methyl group from 5-methyl-THF, or by betaine homocysteine methyltransferase, which utilizes a methyl group from betaine. Thus, folate and betaine can be used interchangeably for the remethylation of homocysteine. Betaine can be obtained through the diet or synthesized from choline. Choline can also be obtained from the diet or it can be synthesized endogenously. Choline may serve as a methyl donor through its conversion to betaine. However, endogenous choline synthesis, like creatine, is another significant consumer of SAM, as roughly 30% is synthesized de novo through a pathway that involves three sequential SAM-dependent methylation reactions catalyzed by phosphatidylethanolamine N-methyltransferase (PEMT) (Vance and Vance, 2004).

Epidemiological studies of choline and As methylation are few, but their results are generally consistent with animal studies. In a cross-sectional analysis (N = 1,016) nested within the Health Effects of Arsenic Longitudinal Study (HEALS), an ongoing cohort in Bangladesh led by Habibul Ahsan, multiple nutrients were found to be associated with As methylation. In multivariate adjusted models, dietary choline intake was positively associated with the secondary methylation index (DMAs/MMAs) (Heck et al., 2007). In a subsequent study of women in Mexico, López-Carrillo et al. 2016 found that higher dietary choline is inversely associated with %MMAs in urine and positively associated with %DMAs and secondary methylation index. However, results from epidemiological studies indicate that dietary betaine has a weaker association with As methylation than choline has. In the HEALS cohort as well as López-Carrillo et al.'s study of Mexican women, betaine intake was not associated with indicators of As methylation capacity; however, both studies relied on estimates of betaine intake from food frequency questionnaires without biological assessment of betaine status.

Our group conducted an 8-week pilot intervention of choline (700 mg/day), betaine (1,000 mg/day) or choline + betaine supplementation in Bangladesh (N =

60). Within-participant changes in %MMAs and %DMAs in urine were significantly different between groups receiving choline, betaine, and choline + betaine ($P < 0.05$) as compared to placebo (M.N. Hall and M.V. Gamble, unpublished data). Although the sample size was small, the data suggest that choline + betaine supplementation resulted in the largest decrease in %MMAs and increase in %DMAs, supporting the hypothesis that both dietary choline and betaine impact As methylation. Enthusiasm for larger intervention studies with choline and betaine has been tempered by the finding that choline supplementation also increased plasma trimethylamine oxide (TMAO). While TMAO has been linked to risk for cardiovascular disease (Tang et al. 2013), it is unclear whether or not this relationship is causal (Zeisel and Warrier, 2017).

2.7 Mechanisms of arsenic toxicity

The biological mechanisms through which As affects cancer and noncancer health outcomes are not fully understood. Arsenic does not directly act as a mutagen in mammal cells. However, it is known that As has multiple modes of action, which may vary between arsenic species. This section will summarize what is known about oxidative stress and protein binding as potential mechanisms of As toxicity. In addition, epigenetic dysregulation will be discussed in detail. The role of epigenetic dysregulation of a mechanism linking chronic As exposure to health outcomes will be further addressed in **Chapters 5-7**.

Oxidative stress

Oxidative stress occurs when a biological system is unable to defend against free radicals or reactive oxygen species (ROS, i.e., oxygen-containing radicals, oxidizing agents, and species readily transformed into radicals) (Betteridge, 2000). Arsenic exposure can generate oxidative stress when As species cycle between oxidation states (Flora,

2011). ROS may in turn act as mutagens (Hei et al., 1998). 8-oxo-29-deoxyguanosine (8-oxodG), a modified base formed by oxygen radicals that may cause G-C to T-A transversions during DNA synthesis (Hayakawa et al., 1995), is a biomarker of oxidative DNA damage. DMAs exposure has been associated with 8-oxodG in rodent studies (Vijayaraghavan et al., 2001; Yamanaka et al., 2001). Although water As concentration was not associated with urinary 8-oxodG in a cohort of adults in Bangladesh (Harper et al., 2014), 8-oxodG concentrations were positively associated with urinary %MMAs in a study of women in Argentina (Engström et al., 2010). Glutathione (GSH), an intercellular antioxidant (Forman et al., 2009) that may protect against As-induced oxidative stress can also serve as a mechanism-based biomarker of oxidative stress (Frijhoff et al., 2015). Water As concentrations were negatively associated with GSH concentrations in blood among As-exposed adults in Bangladesh (Hall et al., 2013).

Protein binding

Arsenic, particularly trivalent species, may interfere with biological processes by binding to cellular proteins. Trivalent As species are able to react with critical sulfhydryl groups of many enzymes. The protein binding of As differs between metabolites due to three coordination sites on $InAs^{III}$, two on $MMAs^{III}$, and only one on $DMAs^{III}$ (Aposhian and Aposhian, 2006). $InAs^{III}$ has been shown to bind to numerous proteins with diverse biological roles including tubulin, poly(ADP-ribose)polymerase (PARP-1), thioredoxin reductase, and estrogen receptor-alpha (Kitchin and Wallace, 2008). $InAs^{III}$ may also disrupt the activity of ten-eleven translocation (TET) enzymes, enzymes involved in active demethyaltion of DNA (discussed in the Epigenetic dysregulation section below). TET enzymes contain zinc finger domains that coordinate with zinc cations, which is important for maintaining the structure of the catalytic site (Hu et al., 2013). $InAs^{III}$ can interact with zinc finger domains, disrupting catalytic activity

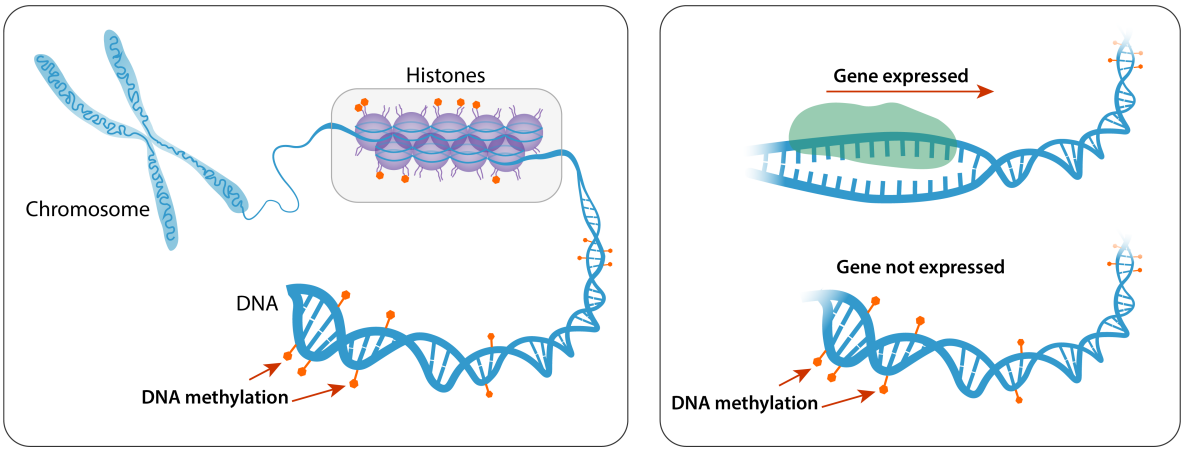


Figure 2.5: Epigenetic modifications: Epigenetic modifications include histone modifications and DNA methylation, which are involved in regulating gene expression throughout life, genomic imprinting, maintaining genomic stability by preventing DNA damage, and establishing chromosomal organization. DNA methylation may control gene transcription by inhibiting transcription factor binding in promoter regions and recruiting proteins that bind to methylated DNA.

(Liu et al., 2015).

Epigenetic dysregulation

Epigenetic modifications are involved in controlling cellular differentiation during development, regulating gene expression throughout life, genomic imprinting, maintaining genomic stability by preventing DNA damage, and establishing chromosomal organization (Jaenisch and Bird, 2003; Putiri and Robertson, 2011). Epigenetic modifications include DNA methylation and histone modifications (Figure 2.5), which influence gene transcription, and microRNAs (miRNAs), which control gene expression posttranscriptionally. Alterations to the epigenome can arise stochastically and as a result of environmental conditions, and contribute to disease onset later in life.

DNA methylation in the form of 5-methylcytosine (5-mC) is perhaps the most commonly studied epigenetic DNA base modification (Smith and Meissner, 2013). DNA methylation is the covalent addition of a CH3 group to the 5-carbon position of cytosine,

and is catalyzed by a family of DNA methyltransferases (DNMT) using the methyl donor SAM (Figure 2.6). DNMT3 alpha (DNMT3A) and beta (DNMT3B) are involved in establishing *de novo* DNA methylation, whereas DNMT1 is involved in maintaining DNA methylation (Okano et al., 1999). DNA methylation most commonly occurs at CpG sites, or a cytosine nucleotide followed by a guanine nucleotide (Figure 2.7). However, DNA methylation can also occur at non-CpG cytosines, at frequencies that vary by cell type; 25% of methylated cytosines have been identified as non-CpG sites in human embryonic stem cells, whereas the proportion is less than 1% in somatic cells (Jang et al., 2017; Laurent et al., 2010; Lister et al., 2009). DNA methylation can control gene expression by inhibiting transcription factor binding in promoter regions and recruiting proteins that bind to methylated DNA (Robertson, 2005) (a schematic drawing transcription regulation is presented in Figure 2.5). Overall, CpG dinucleotides are depleted in the human genome due to mutability of methylated cytosine, however, there are regions containing a relatively high frequency of CpGs, referred to as CpG islands (Gardiner-Garden and Frommer, 1987). The positions of CpGs upstream and downstream of islands are described as North and South shores and shelves, respectively (Figure 2.8). Promotors commonly contain CpG islands, with approximately 70% of promotors having high CpG content (Saxonov et al., 2006). CpGs located in promotor regions are described by their distance from the transcription start site (e.g., TSS 200 indicates CpGs located within 200 base pairs upstream from the transcription start site).

5-hydroxymethylcytosine (5-hmC), an intermediate product of the active demethylation of 5-mC by TET enzymes (Hahn et al. 2014) (Figure 2.6), has also been found to be a stable epigenetic mark (Bachman et al., 2014) associated with gene expression (Ficz et al., 2011; Greco et al., 2016; Jin et al., 2011). 5-hmC is enriched in enhancer regions and gene bodies (Stroud et al., 2011) and methyl-CpG binding

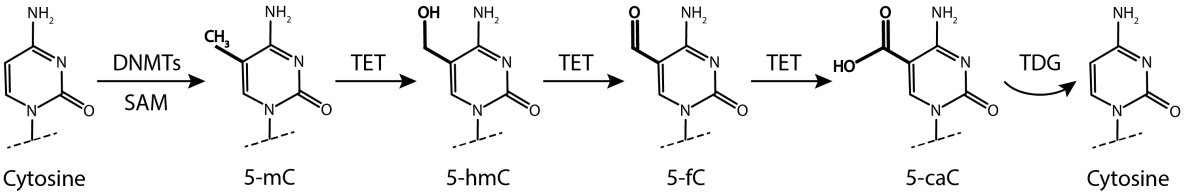


Figure 2.6: DNA methylation occurs through the covalent addition of a -CH_3 group to the 5-carbon position of cytosine in a reaction catalyzed by DNAMT using the methylation donor SAM, forming 5-mC (Jin and Robertson 2013). Demethylation of DNA occurs through TET-mediated oxidation of 5-mC, forming 5-hmC, 5-fC, and 5-caC. 5-caC is subsequently converted to cytosine though TDG-mediated excision repair (Wu and Zhang 2017). Abbreviations: DNMT, DNA methyltransferase; SAM, *S*-Adenosyl methionine; TET, ten-eleven translocation enzymes; 5-mC, 5-methylcytosine; 5-hmC, 5-hydroxymethylcytosine; 5-fC, 5-formylcytosine; 5-caC, 5-carboxylcytosine; TDG, thymine DNA glycosylase.

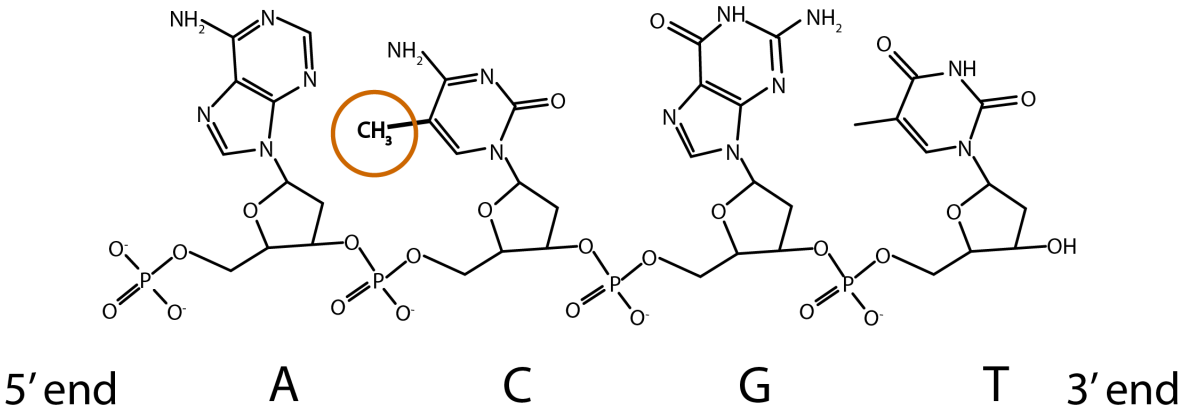


Figure 2.7: DNA methylation at a cytosine-guanine dinucleotide (CpG site).

proteins that interact with 5-hmC have been identified (Mellén et al., 2012; Yildirim et al., 2011).

Changes in global methylation (i.e., the proportion of methylated cytosines in the genome overall) have been associated with disease onset (Wilson et al., 2007) and altered DNAm at CpG sites associated with biologically-relevant genes has been reported in a broad range of conditions (Abi Khalil, 2014; Cribbs et al., 2015; Fradin et al., 2012; Klengel et al., 2014; Lu et al., 2013; Neidhart and Neidhart, 2016; Nilsson et al., 2014). Increasing evidence links environmental exposure to metals, pesticides,

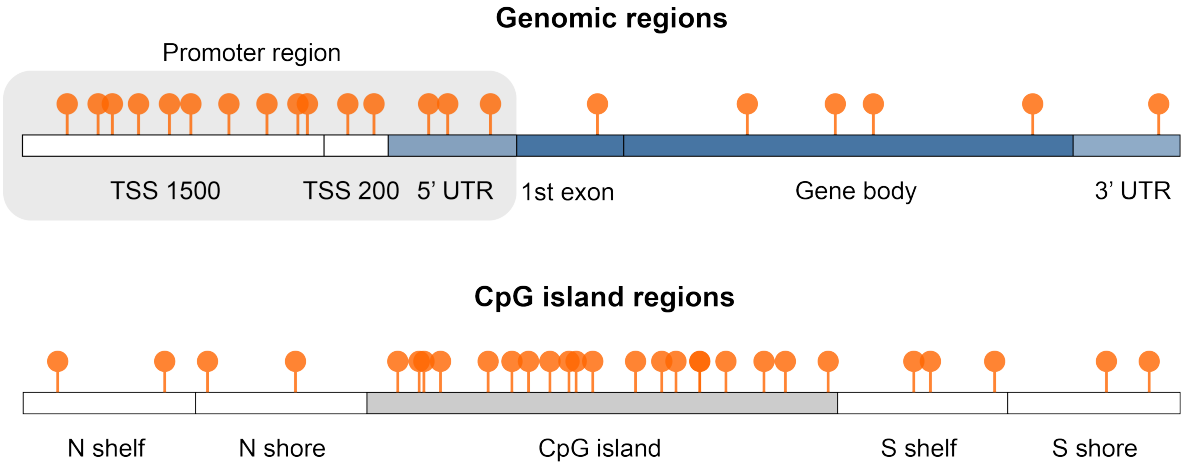


Figure 2.8: Diagram of genomic regions and CpG islands. Orange circles represent DNA methylation. Abbreviations: CpG, cytosine-guanine dinucleotide; TSS, transcription start site; UTR, untranslated region.

air pollution, and synthetic compounds to epigenetic changes detected in blood DNA (Hou et al., 2012).

The observations that exposure to As in both early life (Marshall et al., 2007; Steinmaus et al., 2014) and in adulthood (Steinmaus et al., 2013) increases the risk of diseases later in life suggests that epigenetic dysregulation may be involved in the etiology of diseases associated with As exposure and may contribute to arsenics role as a carcinogen (Bailey et al., 2016; Bailey and Fry, 2014; Carlin et al., 2015). However, the exact nature of epigenetic disruption is not fully understood. Studies have reported conflicting findings regarding the directionality of epigenetic changes induced by As exposure, possibly due to differences between model systems, human populations, dose and duration of As exposure, and measure of DNAm (e.g., global or loci-specific) (Reichard and Puga, 2010).

Our group has investigated the association between As exposure and global DNA methylation. Among adults chronically exposed to As in Bangladesh, urinary, plasma, and total bAs were positively associated with peripheral blood leukocyte DNA methylation using the methyl-incorporation assay, which primarily assess %5-mC; how-

ever, significant associations were only detected among folate-sufficient participants (Pilsner et al., 2007). In a separate cohort, blood and urinary As were positively associated with peripheral blood mononuclear cell DNA methylation regardless of folate status (Niedzwiecki et al., 2013). Although it was hypothesized that As exposure would be negatively associated with DNA methylation due to SAM consumption in As methylation, these positive associations suggest a compensatory epigenetic response to DNA methylation. Furthermore, the association between global DNA methylation and As exposure may be modified by sex. Among participants from both cohorts, there was a trend toward a positive association among males between %5-mC and As exposure as measured by water As, urinary As, and bAs concentrations. In the second cohort, there was a trend toward a positive association among males and a negative association among females between %5-hmC and As exposure (Niedzwiecki et al., 2015).

Inconsistent results regarding the association between As exposure and global DNA methylation have been overserved among other populations. In a Spanish cohort, a negative association was observed between toenail As concentrations and LINE-1 methylation (Tajuddin et al., 2013). In a birth cohort in Bangladesh, high vs. low tercile of maternal urinary As concentration was associated with increased LINE-1 methylation in maternal and cord blood, but no significant associations were observed with the methylation of Alu repeats (Kile et al., 2012).

Associations between As exposure and loci specific DNA methylation have also been investigated, commonly using epigenome-wide association studies (EWAS). Site-specific methylation can be quantified using array hybridization. Illumina (San Diego, CA) has developed high-throughput DNA methylation microarrays (Fan et al., 2006), allowing for efficient profiling of methylation at a single-site resolution in large-scale EWAS. Samples are treated with bisulfite to convert unmethylated cytosines to uracil (Bibikova et al., 2006) (Figure 2.9). Uracil nucleotides are converted to thymine

during PCR, and resultant DNA is hybridized to probes on the Illumina BeadChip array. Methylation status is then determined by a single-base extension that incorporates one of four labeled nucleotides at each CpG site. Labeled nucleotides are fluorescently stained and scanned to determine the intensities of the methylated or unmethylated segments, reflecting the methylation proportion. Figure 2.9 shows a diagram of the bisulfite conversion process and hybridization to type I and type II probes. Methylation is commonly described by Beta-values, which can be interpreted as the percentage of methylation at a given site, or M-values, the logit-transformation of Beta-values (Du et al., 2010). The Illumina Infinium HumanMethylation450 BeadChip (450K), an array that interrogates methylation levels at approximately 485,000 loci representing over 99% of RefSeq genes (Illumina, 2012), has been used in numerous epidemiological studies of the effects of environmental exposures on the methylome in adults, children, cord blood, and human tissues. In 2016, the 450K array was replaced by the Infinium MethylationEPIC BeadChip (850K), covering over 850,000 loci (Illumina, 2015). The 850K array includes 91.1% of sites on the 450K and additional sites located in enhancer and open chromatin regions. A validation study of the 850K demonstrated nearly perfect correlation with the 450K for duplicated sites ($r_{Pearson} = 0.992$) and high replicability ($r_{Pearson} = 0.997$) (Moran et al., 2015).

A 2015 review of EWAS of As exposure identified seven studies (Argos, 2015), however this number has increased substantially in recent years, with 14 published manuscripts identified through a PubMed search in June, 2019. A description of EWAS of As exposure is presented in Appendix Table 2.4. In summary, these studies have investigated associations between *in utero* exposure and DNA methylation measured primarily in cord blood or placenta, and between As exposure in adulthood and DNA methylation measured primarily in whole blood. Nine have measured DNA methylation at birth or in childhood (Bozack et al., 2018a; Broberg et al., 2014; Cardenas et al., 2015;

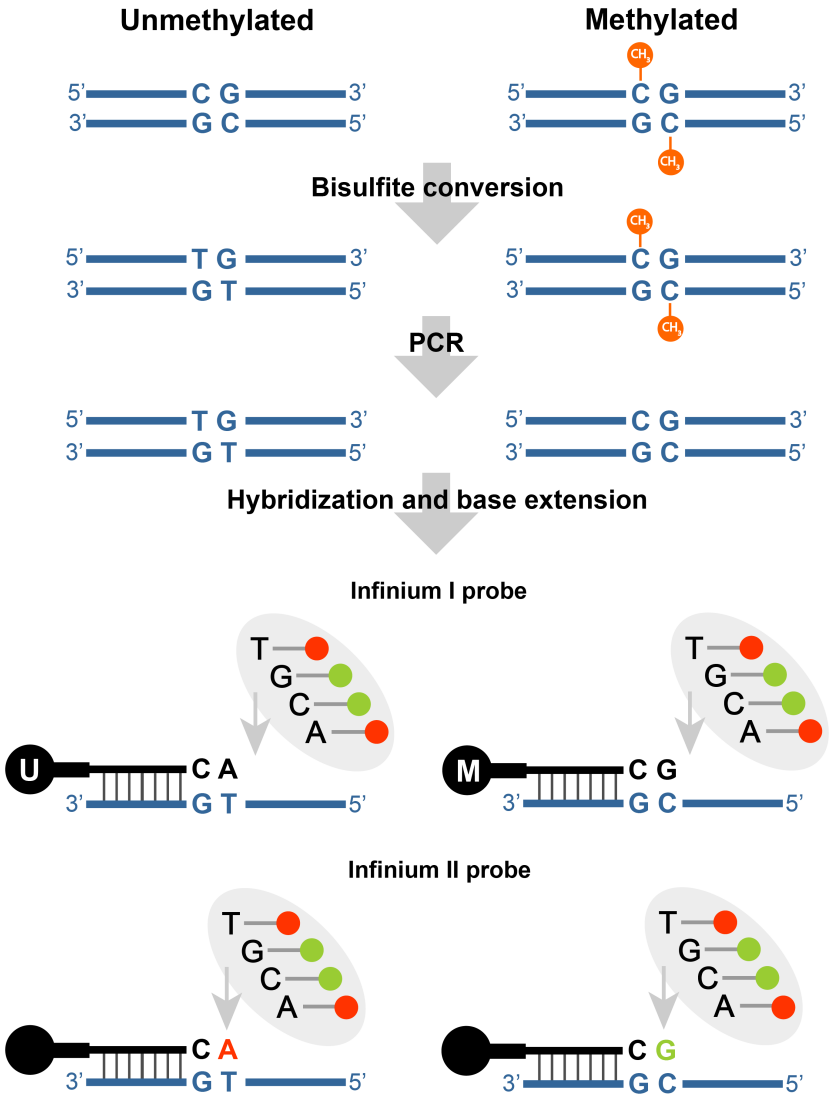


Figure 2.9: Illumina Infinium assay. Samples are treated with bisulfite to convert unmethylated cytosines to uracil, and subsequently uracil nucleotides are converted to thymine nucleotides by PCR. The resultant DNA is hybridized to probes on the Illumina BeadChip array, with hybridization stopping at the locus of interest (type I probes) or one base short of the locus of interest (type II probes). Methylation status is determined by a single-base extension that incorporates one of four DNP- or Biotin-labeled nucleotides. Labeled nucleotides are fluorescently stained and scanned to determine the intensities of the methylated or unmethylated segments, which reflects the methylation proportion of a given locus. The Infinium I assay includes two probes for each loci, targetting either unmethylated or methylated segments. Methylation status is indicated by the nucleotide incorporated downstream of the target cytosine. The Infinium II assay has one probe for each loci, and methylation status is indicated by the guanine or adenine or nucleotide incorporated at the target cytosine.

Gliga et al., 2018; Green et al., 2016; Kaushal et al., 2017; Kile et al., 2014; Koestler et al., 2013; Rojas et al., 2015) and five have included adult populations (Ameer et al., 2017; Argos et al., 2015; Demanelis et al., 2019; Guo et al., 2018; Liu et al., 2014). Most studies have relied on relatively small samples sizes ranging from 44-400, which reduces power, particularly after adjustment for multiple comparisons. The majority of studies have reported findings based on the 450K microarray; Demanelis et al. is the only published EWAS of As exposure to date that has measured DNA methylation using the 850K microarray (Demanelis et al., 2019).

The association between low-level *in utero* As exposure and was assessed in the New Hampshire Birth Cohort. In a subsample of 134 mother-infant pairs, quartiles of maternal urinary As levels at 24-28 weeks gestation were associated with differential methylation levels of 68,353 loci (nominal $P < 0.05$); however no loci remained significant after controlling for multiple comparisons (Koestler et al., 2013). In 343 mother-infant pairs of the same cohort, associations between DNA methylation and placental, postpartum maternal toenail, and maternal urinary As concentrations were assessed. After false discovery rate (FDR) adjustment, 163 CpGs were associated with placental As concentrations ($q < 0.05$), one CpG was associated with maternal toenail As concentrations, and no loci were associated with maternal urinary As concentrations (Green et al., 2016).

Additional EWAS of *in utero* As exposure have focused on medium-high levels of exposure. In a prospective birth cohort in Bangladesh, DNA methylation was measured in 44 mother-infant pairs using the 450k microarray (Kile et al., 2014). Maternal drinking water As concentrations were associated with one CpG in cord blood ($FDR < 0.05$) (Kile et al., 2014), 518 CpGs in placental tissue ($FDR < 0.05$), and no CpGs in human umbilical vein endothelial cells (HUVEC) (Cardenas et al., 2015). In a separate Bangladeshi birth cohort nested in a randomized micronutrient trial, DNA

methylation was measured in cord blood of 127 mother-infant pairs (Brosnan et al., 2011). Maternal urinary As concentrations collected early in the gestational period were associated with methylation at three loci ($FDR < 0.05$) among boys; no significant loci were identified among girls or in relation to maternal urinary As concentrations collected late in gestation. DNA methylation was measured in a cord blood in a subset (N = 38) of Biomarkers of Exposure to Arsenic (BEAR) prospective birth cohort in Mexico (Rojas et al., 2015). Relative to maternal urinary As concentrations, 4,771 differentially methylated loci were identified ($FDR < 0.05$). In a Taiwanese birth cohort, 579 differentially methylation loci were identified in cord blood relative to maternal urinary As ($FDR < 0.05$) (Kaushal et al., 2017)).

In a study conducted by Gliga et al., the association between *in utero* As exposure and DNA methylation at age 9 was assessed in a subset (N = 113) of a birth cohort in Taiwan with moderate levels of As exposure (Gliga et al., 2018). Maternal urinary As levels in early pregnancy were associated with DNA methylation of 9 loci in the full sample, 57 loci among boys, and 15 among girls ($FDR < 0.05$).

Among EWAS of adults, a study in the US with low-level As exposure, no loci were differentially methylated relative to dichotomized toenail As concentrations after correction for multiple comparisons (Liu et al., 2014). Addition EWAS of adults have included moderate-high levels of As exposure. In a study of 96 women in Argentina, six differentially methylated loci were identified relative to urinary As levels (Ameer et al., 2017). In a study of three generations of families in China with and without As exposure (N = 102), drinking water exposure was associated with DNA methylation at 85 loci (Guo et al., 2018).

Two EWAS of As-exposed adults in Bangladesh have been conducted by Habibul Ahsans group. In the Bangladesh Vitamin E and Selenium Trial (BEST), DNA methylation was measured among 400 participants using the 450K microarray (Argos

et al., 2015). Four differentially methylated loci were identified relative to urinary As levels, and three of these loci were also differentially methylated relative to bAs levels at the Bonferroni-correct threshold of $P < 1 \times 10^{-7}$. Most recently, DNA methylation was measured among 396 participants of the Health Effects of Arsenic Longitudinal Study (HEALS) using the 850K microarray (Demanelis et al., 2019). Thirty-four loci were identified as differentially methylated with urinary As levels, and 24 loci were identified as differentially methylated with water As levels ($FDR < 0.05$). Eight CpG were differentially methylated at $FDR < 0.05$ with both measures of exposure.

Across studies, no loci have been consistently found to be differentially methylated in respect to As exposure. However, it is difficult to compare results across studies due to differences in populations, exposure levels, measure of exposure, tissue type, and statistical methods. Argos et al. implemented a lookup approach to assess nominal significance in the BEST study of loci previously identified to be differentially methylated. Eight CpGs previously associated with As exposure or As-related health outcomes in EWAS achieved nominal significance ($P < 0.05$) in BEST (Argos et al., 2015). Demanelis et al. compared 26 top significant loci in HEALS with results from the BEST study (Demanelis et al., 2019). Fifteen were found to be nominally significant in BEST ($P < 0.05$), 14 of which displayed the same direction of association. Comparing results of associated gene ontology analyses may provide information about common pathways affected by As exposure; however, gene ontology analyses have been limited by the small number of As-associated differentially methylated loci identified in most EWAS.

An additional limitation of EWAS of As exposure is limited data regarding changes in gene expression. In the BEAR cohort, Rojas et al. compared methylation levels at loci identified as being differentially methylated with As exposure to expression levels of genes identified as being differentially expressed (Rojas et al., 2015). Among 54 genes exhibiting both differential methylation and expression, methylation and ex-

pression levels were significantly correlated in 12 genes. Among As-exposed women in Argentina, Ameer et al. also measured gene expression (Ameer et al., 2017). However, there was no overlap between genes differentially expressed with As exposure and genes containing As-related differential DNA methylation. Among adults in Bangladesh, Argos et al. assessed expression of the genes containing the top 35 loci associated with As exposure (Argos et al., 2015). Among 28 genes with expression data, DNA methylation was associated with gene expression of 15 genes at $P < 0.10$.

Although studies have consistently linked As exposure to epigenetic dysregulation, additional research is needed to fully understand the impact of As exposure on the epigenome, particularly at low levels of exposure and in diverse populations. Additionally, few studies have addressed the biological implications of observed associations between As exposure and changes in DNA methylation, including the impact on gene expression and health outcomes.

2.8 Summary and Rationale

As described in this chapter, there is strong evidence from observational epidemiological studies of the influence of OCM-related nutrients on As methylation capacity; however, this relationship is not completely understood and current research is primarily based on observations in cross sectional studies. In a previous RCT, our group found that 400 μg/day FA supplementation decreased bAs concentrations and increased As methylation capacity measured in urine compared to placebo among folate deficient adults (Gamble et al., 2006, 2007) (Section 2.6). In FACT, we observed that the decrease in bAs concentrations with 800 μg/day FA exceeded that of placebo among participants recruited independent of folate status. Using data from FACT will allow us to investigate the effects of FA and creatine treatment on As methylation capacity, as well as treatment effect modification by baseline nutritional status of OCM nutrients.

Results from these analyses will be presented in **Chapters 3 and 4**.

It is hypothesized that epigenetic dysregulation may be one biological mechanism underlying the associations between chronic As exposure and health outcomes. Arsenic exposure has been associated with changes in the epigenome, specifically DNA methylation in multiple epidemiological studies (Section 2.7), although results have differed between studies due to differences in populations studied, measures and level of exposure, tissue type, and platform used to measure DNA methylation. In addition, the link between As exposure, epigenetic dysregulation, and health outcomes has not been fully established. Using data from two diverse As-exposed cohorts (the SHS cohort in the US and a prospective birth cohort in Bangladesh) will allow us to investigate the relationship between As exposure, epigenetic dysregulation, and health outcomes. Specifically, we will examine the associations between urinary As levels and loci-specific DNA methylation in SHS, and mediation of the association between *in utero* As exposure and birth outcomes in the Bangladeshi birth cohort. Results from these analyses will be presented in **Chapters 5-7**. Finally, a summary of our conclusions and future directions will be presented in **Chapter 8**.

2.9 Bibliography

Abi Khalil, C. The emerging role of epigenetics in cardiovascular disease. *Therapeutic Advances in Chronic Disease*, 5(4):178–87, 2014. doi: 10.1177/2040622314529325.

Agusa, T; Iwata, H; Fujihara, J; Kunito, T; Takeshita, H; Minh, T B; Trang, P T K; Viet, P H, and Tanabe, S. Genetic polymorphisms in AS3MT and arsenic metabolism in residents of the Red River Delta, Vietnam. *Toxicology and Applied Pharmacology*, 236(2):131–141, 2009. doi: 10.1016/j.taap.2009.01.015.

Ahsan, H; Chen, Y; Kibriya, M G; Slavkovich, V; Parvez, F; Jasmine, F; Gamble, M V, and Graziano, J H. Arsenic metabolism, genetic susceptibility, and risk of premalignant skin lesions in Bangladesh. *Cancer Epidemiology Biomarkers and Prevention*, 16(6):1270–1278, 2007. doi: 10.1158/1055-9965.EPI-06-0676.

Ameer, S S; Engström, K; Hossain, M B; Concha, G; Vahter, M, and Broberg, M. Arsenic exposure from drinking water is associated with decreased gene expression and increased DNA methylation in peripheral blood. *Toxicology and Applied Pharmacology*, 321:57–66, 2017. doi: 10.1016/J.TAAP.2017.02.019.

Aposhian, H Vasken and Aposhian, M M. Arsenic toxicology: five questions. *Chemical Research in Toxicology*, 19(1):1–15, 2006. doi: 10.1021/tx050106d.

Argos, M. Arsenic exposure and epigenetic alterations: recent findings based on the Illumina 450K DNA methylation array. *Current Environmental Health Reports*, 2(2): 137–44, 2015. doi: 10.1007/s40572-015-0052-1.

Argos, M; Chen, L; Jasmine, F; Tong, L; Pierce, B L; Roy, S; Paul-Brutus, R; Gamble, M V; Harper, K N; Parvez, F; Rahman, M; Rakibuz-Zaman, M; Slavkovich, V; Baron, J A; Graziano, J H; Kibriya, M G, and Ahsan, H. Gene-specific differential DNA methylation and chronic arsenic exposure in an epigenome-wide association study of adults in Bangladesh. *Environmental Health Perspectives*, 123(1):64–71, 2015. doi: 10.1289/ehp.1307884.

Ayotte, J D; Gronberg, J-A M, and Apodaca, L E. Trace Elements and radon in groundwater across the United States, 1992-2003. Technical report, U.S. Geological Survey Scientific Investigations Report. URL `https://water.usgs.gov/nawqa/trace/pubs/sir2011-5059/`.

Ayotte, J D; Medalie, L; Qi, S L; Backer, L C, and Nolan, B T. Estimating the high-arsenic domestic-well population in the conterminous United States. *Environmental Science and Technology*, 51(21):12443–12454, 2017. doi: 10.1021/acs.est.7b02881.

Bachman, M; Uribe-Lewis, S; Yang, X; Williams, M; Murrell, A, and Balasubramanian, S. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nature Chemistry*, 6(12):1049–1055, 2014. doi: 10.1038/nchem.2064.

Bailey, K A and Fry, R C. Arsenic-associated changes to the epigenome: what are the functional consequences? *Current Environmental Health Reports*, pages 22–34, 2014. doi: 10.1007/s40572-013-0002-8.

Bailey, K A; Smith, A H; Tokar, E J; Graziano, J H; Kim, K-W; Navasumrit, P; Ruchirawat, M; Thiantanawat, A; Suk, W A, and Fry, R C. Mechanisms underlying latent disease risk associated with early-life arsenic exposure: current research trends and scientific gaps. *Environmental Health Perspectives*, 124(2):170–175, 2016. doi: 10.1289/ehp.1409360.

Balakrishnan, P; Vaidya, D; Franceschini, N; Voruganti, V S; Gribble, M O; Haack, K; Laston, S; Umans, J G; Francesconi, K A; Goessler, W; North, K E; Lee, E; Yracheta, J; Best, L G; MacCluer, J W; Kent, J; Cole, S A, and Navas-Acien, A. Association of cardiometabolic genes with arsenic metabolism biomarkers in American Indian communities: the Strong Heart Family Study (SHFS). *Environmental Health Perspectives*, 125(1):15–22, 2017. doi: 10.1289/EHP251.

Bangladesh Bureau of Statistics and United Nation Children's Fund. Bangladesh multiple indicator cluster survey 2012-2013 final report. Technical report, Dhaka, Bangladesh, 2015.

Basu, A; Mitra, S; Chung, J; Guha Mazumder, DN; Ghosh, N; Kalman, D; von Ehrenstein, O S; Steinmaus, C; Liaw, J, and Smith, A H. Creatinine, diet, micronutrients, and arsenic methylation in West Bengal, India. *Environmental Health Perspectives*, 119(9):1308–1313, 2011. doi: 10.1289/ehp.1003393.

Betteridge, D J. What is oxidative stress? *Metabolism*, 49(2):3–8, 2000. doi: 10.1016/S0026-0495(00)80077-3.

Bibikova, M; Lin, Z; Zhou, L; Chudin, E; Garcia, Eliza W; Wu, B; Doucet, D; Thomas, N J; Wang, Y; Vollmer, E; Goldmann, T; Seifart, C; Jiang, W; Barker, D L; Chee, M S; Floros, J, and Fan, J-B. High-throughput DNA methylation profiling using universal bead arrays. *Genome Research*, 16(3):383–93, 2006. doi: 10.1101/gr.4410706.

Bloom, M S; Surdu, S; Neamtiu, I A, and Gurzau, E S. Maternal arsenic exposure and birth outcomes: a comprehensive review of the epidemiologic literature focused on drinking water HHS Public Access. *International Journal of Hygiene and Environmental Health*, 217(7):709–719, 2014. doi: 10.1016/j.ijheh.2014.03.004.

Bloom, M S; Neamtiu, I A; Surdu, S; Pop, C; Anastasiu, D; Appleton, A A; Fitzgerald, E F, and Gurzau, E S. Low level arsenic contaminated water consumption and birth outcomes in Romania - an exploratory study. *Reproductive Toxicology*, 59:8–16, 2016. doi: 10.1016/j.reprotox.2015.10.012.

Bozack, A K; Cardenas, A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Christiani, D C, and Kile, M L. DNA methylation in cord blood as mediator of the

association between prenatal arsenic exposure and gestational age. *Epigenetics*, 13 (9):923–940, 2018a. doi: 10.1080/15592294.2018.1516453.

Bozack, A K; Hall, M N; Liu, X; Ilievski, V; Lomax-Luu, A M; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T, and Graziano, J H. Folic acid supplementation enhances arsenic methylation: results from a folic acid and creatine supplementation randomized controlled trial in Bangladesh. *American Journal of Clinical Nutrition*, 109(2):380–391, 2018b. doi: 10.1093/ajcn/nqy148.

Bozack, A K; Saxena, R, and Gamble, M V. Nutritional influences on one-carbon metabolism: effects on arsenic methylation and toxicity. *Annual Review of Nutrition*, 38:401–429, 2018c. doi: 10.1146/annurev-nutr-082117-051757.

Broberg, K; Ahmed, S; Engström, K; Hossain, M B; Jurkovic Mlakar, S; Bottai, M; Grandér, M; Raqib, R, and Vahter, M. Arsenic exposure in early pregnancy alters genome-wide DNA methylation in cord blood, particularly in boys. *Journal of Developmental Origins of Health and Disease*, 5(4):288–98, 2014. doi: 10.1017/S2040174414000221.

Brosnan, J T; da Silva, R P, and Brosnan, M E. The metabolic burden of creatine synthesis. *Amino Acids*, 40(5):1325–31, 2011. doi: 10.1007/s00726-011-0853-y.

Brouwer, O F F; Onkenhout, W; Edelbroek, P M; de Kom, J F M; de Wolff, F A; Peters, A C B; Edelbroed, P M; de Kom, J F M, and de Wolff, F A. Increased neurotoxicity of arsenic in methylenetetrahydrofolate reductase deficiency. *Clinical Neurology and Neurosurgery*, 94(4):307–310, 1992. doi: 10.1016/0303-8467(92)90179-7.

Buchet, J P; Lauwerys, R, and Roels, H. Comparison of the urinary excretion of arsenic metabolites after a single oral dose of sodium arsenite, monomethylarsonate, or dimethylarsinate in man. *International Archives of Occupational and Enironmental Health*, 48(1):71–9, 1981.

Cardenas, A; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Wright, R O; Christiani, D C, and Kile, M L. *In utero* arsenic exposure and epigenome-wide associations in placenta, umbilical artery, and human umbilical vein endothelial cells. *Epigenetics*, 10(11):1054–1063, 2015. doi: 10.1080/15592294.2015.1105424.

Carlin, D J; Naujokas, M F; Bradham, K D; Cowden, J; Heacock, M; Henry, H F; Lee, J S; Thomas, D J; Thompson, C; Tokar, E J; Waalkes, M P; Birnbaum, L S, and Suk, W A. Arsenic and environmental health: state of the science and future research opportunities. *Environmental Health Perspectives*, 2015. doi: 10.1289/ehp.1510209.

Carreras, C W and Santi, D V. The catalytic mechanism and structure of thymidylate synthase. *Annual Review of Biochemistry*, 64(1):721–762, 1995. doi: 10.1146/annurev.bi.64.070195.003445.

Challenger, F. Biological methylation. *Chemical Reviews*, 36(3):315–361, 1945. doi: 10.1021/cr60115a003.

Chen, J-W; Wang, S-L; Wang, Y-H; Sun, C-W; Huang, Y-L; Chen, C-J, and Li, W-F. Arsenic methylation, GSTO1 polymorphisms, and metabolic syndrome in an arseniasis endemic area of southwestern Taiwan. *Chemosphere*, 88(4):432–8, 2012. doi: 10.1016/j.chemosphere.2012.02.059.

Chen, Y; Han, Y-H; Cao, Y; Zhu, Y-G; Rathinasabapathi, B, and Ma, L Q. Arsenic transport in rice and biological solutions to reduce arsenic risk from rice. *Frontiers in Plant Science*, 8:268, 2017. doi: 10.3389/fpls.2017.00268.

Chen, Y-C; Guo, Y-L L; Su, H-J J; Hsueh, Y-M; Smith, T J; Ryan, L M; Lee, M-S; Chao, S-C; Lee, J Y-Y, and Christiani, D C. Arsenic methylation and skin cancer risk in southwestern Taiwan. *Journal of Occupational and Environmental Medicine*, 45(3):241–8, 2003a.

Chen, Y-C; Su, H-J J; Guo, Y-L L; Hsueh, Y-M; Smith, T J; Ryan, L M; Lee, M-S, and Christiani, D C. Arsenic methylation and bladder cancer risk in Taiwan. *Cancer Causes and Control*, 14(4):303–10, 2003b. doi: 10.1023/A:1023905900171.

Chowdhury, U K; Biswas, B K; Chowdhury, T R; Samanta, G; Mandal, B K; Basu, G C; Chanda, C R; Lodh, D; Saha, K C; Mukherjee, S K; Roy, Sibtosh; Kabir, S; Quamruzzaman, Q, and Chakraborti, D. Groundwater arsenic contamination in Bangladesh and West Bengal, India. *Environ Health Perspect*, 108(5):393–7, 2000.

Chung, J S; Kalman, D A; Moore, L E; Kosnett, M J; Arroyo, A P; Beeris, M; Mazumder, D N G; Hernandez, A L, and Smith, A H. Family correlations of arsenic methylation patterns in children and parents exposed to high concentrations of arsenic in drinking water. *Environmental Health Perspectives*, 110(7):729–33, 2002.

Concha, G; Vogler, G; Lezcano, D; Nermell, B, and Vahter, M. Exposure to inorganic arsenic metabolites during early human development. *Toxicological Sciences*, 44(2): 185–190, 1998.

Cribbs, A; Feldmann, M, and Oppermann, U. Towards an understanding of the role of DNA methylation in rheumatoid arthritis: therapeutic and diagnostic implications. *Therapeutic Advances in Musculoskeletal Disease*, 7(5):206–19, 2015. doi: 10.1177/1759720X15598307.

Del Razo, L M; Styblo, M; Cullen, W R, and Thomas, D J. Determination of trivalent methylated arsenicals in biological matrices. *Toxicology and Applied Pharmacology*, 174(3):282–293, 2001. doi: 10.1006/taap.2001.9226.

Demanelis, K; Argos, M; Tong, L; Shinkle, J; Sabarinathan, M; Rakibuz-Zaman, M; Sarwar, G; Shahriar, H; Islam, T; Rahman, M; Yunus, M; Graziano, J H.; Broberg, K; Engström, K; Jasmine, F; Ahsan, H, and Pierce, B L. Association of arsenic exposure

with whole blood DNA methylation: an epigenome-wide study of Bangladeshi adults. *Environmental Health Perspectives*, 127(5), 2019. doi: 10.1289/EHP3849.

Deminice, R; Portari, G V; Vannucchi, H, and Jordao, A A. Effects of creatine supplementation on homocysteine levels and lipid peroxidation in rats. *British Journal of Nutrition*, 102(1):110–116, 2008. doi: 10.1017/S0007114508162985.

Drobna, Z; Naranmandura, H; Kubachka, K M; Edwards, B C; Herbin-Davis, K; Styblo, M; Le, X Chris; Creed, J T; Maeda, N; Hughes, M F, and Thomas, D J. Disruption of the arsenic (+3 oxidation state) methyltransferase gene in the mouse alters the phenotype for methylation of arsenic and affects distribution and retention of orally administered arsenate. *Chemical Research in Toxicology*, 22(10):1713–1720, 2009. doi: 10.1021/tx900179r.

Du, P; Zhang, X; Huang, C-C; Jafari, N; Kibbe, W A; Hou, L, and Lin, S M. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, 11:587, 2010. doi: 10.1186/1471-2105-11-587.

Engström, K; Vahter, M; Mlakar, S J; Concha, G; Nermell, B; Raqib, R; Cardozo, A, and Broberg, K. Polymorphisms in arsenic(+III oxidation state) methyltransferase (AS3MT) predict gene expression of AS3MT as well as arsenic metabolism. *Environmental Health Perspectives*, 119(2):182–188, 2011. doi: 10.1289/ehp.1002471.

Engström, K S; Vahter, M; Lindh, C; Teichert, F; Singh, R; Concha, G; Nermell, B; Farmer, P B; Strömberg, U, and Broberg, K. Low 8-oxo-7,8-dihydro-2-deoxyguanosine levels and influence of genetic background in an Andean population exposed to high levels of arsenic. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 683(1-2):98–105, 2010. doi: 10.1016/j.mrfmmm.2009.10.014.

Engström, K S; Hossain, M B; Lauss, M; Ahmed, S; Raqib, R; Vahter, M, and Broberg, K. Efficient arsenic metabolism -the AS3MT haplotype is associated with DNA methylation and expression of multiple genes around AS3MT. *PLoS ONE*, 8 (1):e53732, 2013. doi: 10.1371/journal.pone.0053732.

Engström, K S; Vahter, M; Fletcher, T; Leonardi, G; Goessler, W; Gurzau, E; Koppova, K; Rudnai, P; Kumar, R, and Broberg, K. Genetic variation in arsenic (13 Oxidation State) methyltransferase (*AS3MT*), arsenic metabolism and risk of basal cell carcinoma in a European population. *Environmental and Molecular Mutagenesis*, 56(1):60–69, 2015. doi: 10.1002/em.21896.

Fan, JB; Gunderson, K L; Bibikova, M; Yeakley, J M; Chen, J; Wickham Garcia, E; Lebruska, L L; Laurent, M; Shen, R, and Barker, D. Illumina universal bead arrays. *Methods in Enzymology*, 410:57–73, 2006. doi: 10.1016/S0076-6879(06)10003-8.

Ferm, V H and Hanlon, D P. Constant rate exposure of pregnant hamsters to arsenate during early gestation. *Environmental Research*, 37(2):425–32, 1985.

Ficz, G; Branco, M R; Seisenberger, S; Santos, F; Krueger, F; Hore, T A; Marques, C J; Andrews, S, and Reik, W. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*, 473(7347):398–402, 2011. doi: 10.1038/nature10008.

Flanagan, S V; Johnston, R B, and Zheng, Y. Arsenic in tube well water in Bangladesh: health and economic impacts and implications for arsenic mitigation. *Bulletin of the World Health Organization*, 90(September):839–46, 2012. doi: 10.2471/BLT.11.101253.

Fleischer, M. Recent estimates of the abundances of the elements in the earth's crust, Geological Survey Circular 285. Technical report, United States Department of the Interior, Geological Survey, Washington, DC, 1953. URL `https://pubs.usgs.gov/circ/1953/0285/report.pdf`.

Flora, C J S. Arsenic-induced oxidative stress and its reversibility. *Free Radical Biology and Medicine*, 51(2):257–281, 2011. doi: 10.1016/j.freeradbiomed.2011.04.008.

Food and Drug Administration. Arsenic in rice and rice products: risk assessment report. Technical report, Center for Food Safety and Applied Nutrition, US Department of Health and Human Services, College Park, MD, 2016. URL `http://www.fda.gov/Food/FoodScienceResearch/RiskSafetyAssessment/default.htm`.

Food and Drug Administration. Total Diet Study: elements results summary statisticsmarket baskets 2006 through 2013. Technical report, Center for Food Safety and Applied Nutrition, US Department of Health and Human Services, College Park, MD, 2017.

Forman, H J; Zhang, H, and Rinna, A. Glutathione: Overview of its protective roles, measurement, and biosynthesis. *Molecular Aspects of Medicine*, 30(1-2):1–12, 2009. doi: 10.1016/j.mam.2008.08.006.

Fradin, D; Le Fur, S; Mille, C; Naoui, N; Groves, C; Zelenika, D; McCarthy, M I; Lathrop, M, and Bougnères, P. Association of the CpG methylation pattern of the proximal insulin gene promoter with type 1 diabetes. *PloS ONE*, 7(5):e36278, 2012. doi: 10.1371/journal.pone.0036278.

Freire, C; Amaya, E; Gil, F; Murcia, M; LLop, S; Casas, M; Vrijheid, M; Lertxundi, A; Irizar, A; Fernández-Tardón, G; Castro-Delgado, R V; Olea, N; Fernández, M F, and INMA Project. Placental metal concentrations and birth outcomes: the Environment and Childhood (INMA) project. *International Journal of Hygiene and Environmental Health*, 222(3):468–478, 2019. doi: 10.1016/j.ijheh.2018.12.014.

Frijhoff, J; Winyard, P G; Zarkovic, N; Davies, S S; Stocker, R; Cheng, D; Knight, A R; Taylor, E L; Oettrich, J; Ruskovska, T; Gasparovic, A C; Cuadrado, A; Weber, D; Poulsen, H E; Grune, T; Schmidt, H H H W, and Ghezzi, P. Clinical relevance of biomarkers of oxidative stress. *Antioxidants and Redox Signaling*, 23(14):1144–70, 2015. doi: 10.1089/ars.2015.6317.

Frith, J. Arsenic the poison of kings and the saviour of syphilis. *Journal of Military and Veterans' Health*, 21(4), 2013.

Gamble, M V; Ahsan, H; Liu, X; Factor-Litvak, P; Ilievski, V; Slavkovich, VV; Parvez, F, and Graziano, J H. Folate and cobalamin deficiencies and hyperhomocysteinemia in Bangladesh. *American Journal of Clinical Nutrition*, 81(6):1372–1377, 2005a.

Gamble, M V; Liu, X; Ahsan, H; Pilsner, R; Ilievski, V; Slavkovich, V; Parvez, F; Levy, D; Factor-Litvak, P, and Graziano, J H. Folate, homocysteine, and arsenic metabolism in arsenic-exposed individuals in Bangladesh. *Environmental Health Perspectives*, 113(12):1683–8, 2005b. doi: 10.1289/ehp.8084.

Gamble, M V; Liu, X; Ahsan, H; Pilsner, J R; Ilievski, V; Slavkovich, V; Parvez, F; Chen, Y; Levy, D; Factor-Litvak, P, and Graziano, J H. Folate and arsenic metabolism: a double-blind, placebo-controlled folic acidsupplementation trial in Bangladesh. *American Journal of Clinical Nutrition*, 84(5):1093–101, 2006. doi: 10.1093/ajcn/84.5.1093.

Gamble, M V; Liu, X; Slavkovich, V; Pilsner, J R; Ilievski, V; Factor-Litvak, P; Levy, D; Alam, S; Islam, M; Parvez, F; Ahsan, H, and Graziano, J H. Folic acid supplementation lowers blood arsenic. *American Journal of Clinical Nutrition*, 86(4):1202–9, 2007. doi: 10.1093/ajcn/86.4.1202.

Gao, J; Tong, L; Argos, M; Bryan, M S; Ahmed, A; Rakibuz-Zaman, M; Kibriya, M G; Jasmine, F; Slavkovich, V; Graziano, J H; Ahsan, H, and Pierce, B L. The genetic architecture of arsenic metabolism efficiency: a SNP-based heritability study of Bangladeshi adults. *Environmental Health Perspectives*, 123(10):985–92, 2015. doi: 10.1289/ehp.1408909.

Gardiner-Garden, M and Frommer, M. CpG Islands in vertebrate genomes. *Journal of Molecular Biology*, 196(2):261–282, 1987. doi: 10.1016/0022-2836(87)90689-9.

Gardner, R M; Nermell, B; Kippler, M; Grandér, M; Li, L; Ekström, E-C; Rahman, A; Lönnerdal, B; Hoque, A M W, and Vahter, M. Arsenic methylation efficiency increases during the first trimester of pregnancy independent of folate status. *Reproductive Toxicology*, 31(2):210–218, 2011.

Gilbert-Diamond, D; Li, Z; Perry, A E; Spencer, S K; Gandolfi, A J, and Karagas, M R. A population-based casecontrol study of urinary arsenic species and squamous cell carcinoma in New Hampshire, USA. *Environmental Health Perspectives*, 121(10): 1154–1160, 2013. doi: 10.1289/ehp.1206178.

Gilbert-Diamond, D; Emond, J A; Baker, E R; Korrick, S A, and Karagas, M R. Relation between *in utero* arsenic exposure and birth outcomes in a cohort of mothers and their newborns from New Hampshire. *Environmental Health Perspectives*, 124 (8):1299–307, 2016. doi: 10.1289/ehp.1510065.

Gliga, A R; Engström, K; Kippler, M; Skröder, H; Ahmed, S; Vahter, M; Raqib, R, and Broberg, K. Prenatal arsenic exposure is associated with increased plasma IGFBP3 concentrations in 9-year-old children partly via changes in DNA methylation. *Archives of Toxicology*, 92(8):2487–2500, 2018. doi: 10.1007/s00204-018-2239-3.

Goodrich, J M; Ingle, M E; Domino, S E; Treadwell, M C; Dolinoy, D C; Burant, C; Meeker, J D, and Padmanabhan, V. First trimester maternal exposures to endocrine disrupting chemicals and metals and fetal size in the Michigan Mother-Infant Pairs study. *Journal of Developmental Origins of Health and Disease*, pages 1–12, 2019. doi: 10.1017/S204017441800106X.

Govarts, E; Remy, S; Bruckers, L; Den Hond, E; Sioen, I; Nelen, V; Baeyens, W; Nawrot, T; Loots, I; Van Larebeke, N, and Schoeters, G. Combined effects of prenatal exposures to environmental chemicals on birth weight. *International Journal of Environmental Research and Public Health*, 13(5):495, 2016. doi: 10.3390/ijerph13050495.

Greco, C M; Kunderfranco, P; Rubino, M; Larcher, V; Carullo, P; Anselmo, Z; Kurz, K; Carell, T; Angius, A; Latronico, M V G; Papait, R, and Condorelli, G. DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. *Nature Communications*, 7:12418, 2016. doi: 10.1038/ncomms12418.

Green, B B; Karagas, M R; Punshon, T; Jackson, B P; Robbins, D J; Houseman, E A, and Marsit, C J. Epigenome-wide assessment of DNA methylation in the placenta and arsenic exposure in the New Hampshire Birth Cohort Study. *Environmental Health Perspectives*, 124(8):1253–60, 2016. doi: 10.1289/ehp.1510437.

Guan, H; Piao, F; Zhang, X; Li, X; Li, Q; Xu, L; Kitamura, F, and Yokoyama, K. Prenatal exposure to arsenic and its effects on fetal development in the general population of Dalian. *Biological Trace Element Research*, 149(1):10–15, 2012. doi: 10.1007/s12011-012-9396-7.

Guo, X; Chen, X; Wang, J; Liu, Z; Gaile, D; Wu, H; Yu, G; Mao, G; Yang, Z; Di, Z; Guo, X; Cao, L; Chang, P; Kang, B; Chen, J; Gao, W, and Ren, X. Multi-generational impacts of arsenic exposure on genome-wide DNA methylation and the implications for arsenic-induced skin lesions. *Environment International*, 119:250–263, 2018. doi: 10.1016/J.ENVINT.2018.06.024.

Guthmiller, P; Van Pilsum, J F; Boen, J R, and McGuire, D M. Cloning and sequencing of rat kidney L-arginine:glycine amidinotransferase. Studies on the mechanism of regulation by growth hormone and creatine. *Journal of Biological Chemistry*, 269 (26):17556–60, 1994.

Hall, M; Gamble, M; Slavkovich, V; Liu, X; Levy, D; Cheng, Z; van Geen, A; Yunus, M; Rahman, M; Pilsner, J R, and Graziano, J. Determinants of arsenic metabolism: blood arsenic metabolites, plasma folate, cobalamin, and homocysteine concentrations in maternalnewborn pairs. *Environmental Health Perspectives*, 115(10):1503–1509, 2007. doi: 10.1289/ehp.9906.

Hall, M N; Liu, X; Slavkovich, V; Ilievski, V; Pilsner, J R; Alam, S; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate, cobalamin, cysteine, homocysteine, and arsenic metabolism among children in Bangladesh. *Environmental Health Perspectives*, 117(5), 2009. doi: 10.1289/ehp.0800164.

Hall, M N; Niedzwiecki, M; Liu, X; Harper, K N; Alam, S; Slavkovich, V; Ilievski, V; Levy, D; Siddique, A B; Parvez, F; Mey, J L; van Geen, A; Graziano, J, and Gamble, M V. Chronic arsenic exposure and blood glutathione and glutathione disulfide concentrations in Bangladeshi adults. *Environmental Health Perspectives*, 121(9): 1068–74, 2013. doi: 10.1289/ehp.1205727.

Hansen, H R.; Raab, A; Jaspars, M; Milne, B F, and Feldmann, J. Sulfur-containing arsenical mistaken for dimethylarsinous acid [DMA(III)] and identified as a natural metabolite in urine: major implications for studies on arsenic metabolism and toxicity. *Chemical Research in Toxicology*, 17(8):1086–1091, 2004. doi: 10.1021/tx049978q.

Harper, K N; Liu, X; Hall, M N; Ilievski, V; Oka, J; Calancie, L; Slavkovich, V; Levy, D; Siddique, A; Alam, S; Mey, J L; van Geen, A; Graziano, J H, and Gamble, M V. A dose-response study of arsenic exposure and markers of oxidative damage in Bangladesh. *Journal of Occupational and Environmental Medicine*, 56(6):652–8, 2014. doi: 10.1097/JOM.0000000000000166.

Hayakawa, H; Taketomi, A; Sakumi, K; Kuwano, M, and Sekiguchi, M. Generation and elimination of 8-oxo-7,8-dihydro-2'-deoxyguanosine 5'-triphosphate, a mutagenic substrate for DNA synthesis, in human cells. *Biochemistry*, 34(1):89–95, 1995. doi: 10.1021/bi00001a011.

Hayakawa, T; Kobayashi, Y; Cui, X, and Hirano, S. A new metabolic pathway of arsenite: arsenic-glutathione complexes are substrates for human arsenic methyltransferase Cyt19. *Archives of Toxicology*, 79(4):183–191, 2005. doi: 10.1007/s00204-004-0620-x.

Heck, J E; Gamble, M V; Chen, Y; Graziano, J H; Slavkovich, V; Parvez, F; Baron, J A; Howe, G R, and Ahsan, H. Consumption of folate-related nutrients and metabolism of arsenic in Bangladesh. *American Journal of Clinical Nutrition*, 85(5):1367–74, 2007. doi: 10.1093/ajcn/85.5.1367.

Hei, T K; Liu, S X, and Waldren, C. Mutagenicity of arsenic in mammalian cells: role of reactive oxygen species. *PNAS*, 95(14):8103–7, 1998. doi: 10.1073/pnas.95.14.8103.

Hernández, A; Xamena, N; Sekaran, C; Tokunaga, H; Sampayo-Reyes, A; Quinteros, D; Creus, A, and Marcos, R. High arsenic metabolic efficiency in AS3MT287Thr allele carriers. *Pharmacogenetics and Genomics*, 18(4):349–55, 2008a. doi: 10.1097/FPC.0b013e3282f7f46b.

Hernández, A; Xamena, N; Surrallés, J; Sekaran, C; Tokunaga, H; Quinteros, D; Creus, A, and Marcos, R. Role of the Met 287 Thr polymorphism in the AS3MT gene on the metabolic arsenic profile. *Mutation Research*, 637:80–92, 2008b. doi: 10.1016/j.mrfmmm.2007.07.004.

Hill, D S; Wlodarczyk, B J, and Finnell, R H. Reproductive consequences of oral arsenate exposure during pregnancy in a mouse model. *Birth Defects Research (Part B)*, 83:40–7, 2008. doi: 10.1002/bdrb.20142.

Hood, R D and Bishop, S L. Teratogenic effects of sodium arsenate in mice. *Archives of Environmental Health*, 24(1):62–65, 1972. doi: 10.1080/00039896.1972.10666051.

Hood, Ronald D. Effects of sodium arsenite on fetal development. *Bulletin of Environmental Contamination and Toxicology*, 7(4):216–222, 1972. doi: 10.1007/BF01684401.

Hopenhayn, C; Ferreccio, C; Browning, S R.; Huang, B; Peralta, C; Gibb, H, and Hertz-Picciotto, I. Arsenic exposure from drinking water and birth weight. *Epidemiology*, 14(5):593–602, 2003. doi: 10.1097/01.ede.0000072104.65240.69.

Hou, L; Zhang, X; W, D, and Baccarelli, A. Environmental chemical exposures and human epigenetics. *International Journal of Epidemiology*, 41(1):79–105, 2012. doi: 10.1093/ije/dyr154.

Hsueh, Y M; Chiou, H Y; Huang, Y L; Wu, W L; Huang, C C; Yang, M H; Lue, L C; Chen, G S, and Chen, C J. Serum beta-carotene level, arsenic methylation capability, and incidence of skin cancer. *Cancer Epidemiology and Prevention Biomarkers*, 6(8): 589–96, 1997.

Hu, L; Li, Z; Cheng, J; Rao, Q; Gong, W; Liu, M; Shi, Y G; Zhu, J; Wang, P, and Xu, Y. Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell*, 155(7):1545–1555, 2013. doi: 10.1016/j.cell.2013.11.020.

Hu, X; Zheng, T; Cheng, Y; Holford, T; Lin, S; Leaderer, B; Qiu, J; Bassig, B A; Shi, K; Zhang, Y; Niu, J; Zhu, Y; Li, Y; Guo, H; Chen, Q; Zhang, J; Xu, S, and Jin, Y. Distributions of heavy metals in maternal and cord blood and the association with infant birth weight in China. *The Journal of Reproductive Medicine*, 60(1-2): 21–9, 2015.

Huang, Y-K; Tseng, C-H; Huang, Y-L; Yang, M-H; Chen, C-J, and Hsueh, Y-M. Arsenic methylation capability and hypertension risk in subjects living in arseniasis-hyperendemic areas in southwestern Taiwan. *Toxicology and Applied Pharmacology*, 218:135–142, 2006. doi: 10.1016/j.taap.2006.10.022.

Huang, Y-K; Huang, Y-L; Hsueh, Y-M; Yang, M-H; Wu, M-M; Chen, S-Y; Hsu, L-I, and Chen, C-J. Arsenic exposure, urinary arsenic speciation, and the incidence of urothelial carcinoma: a twelve-year follow-up study. *Cancer Causes and Control*, 19 (8):829–839, 2008a. doi: 10.1007/s10552-008-9146-5.

Huang, Y-K; Pu, Y-S; Chung, C-J; Shiue, H-S; Yang, M-H; Chen, C-J, and Hsueh, Y-M. Plasma folate level, urinary arsenic methylation profiles, and urothelial carcinoma susceptibility. *Food and Chemical Toxicology*, 46(3):929–938, 2008b. doi: 10.1016/j.fct.2007.10.017.

Hughes, M F; Edwards, B C; Herbin-Davis, K M; Saunders, J; Styblo, M, and Thomas, D J. Arsenic (+3 oxidation state) methyltransferase genotype affects steady-state distribution and clearance of arsenic in arsenate-treated mice. *Toxicology and Applied Pharmacology*, 249(3):217–223, 2010. doi: 10.1016/j.taap.2010.09.017.

Huyck, K L; Kile, M L; Mahiuddin, G; Quamruzzaman, Q; Rahman, M; Breton, C V; Dobson, C B; Frelich, J; Hoffman, E; Yousuf, J; Afroz, S; Islam, S, and Christiani, D C. Maternal arsenic exposure associated with low birth weight in Bangladesh. *Journal of Occupational and Environmental Medicine*, 49(10):1097–1104, 2007. doi: 10.1097/JOM.0b013e3181566ba0.

IARC Working Group. Arsenic, metals, fibres, and dusts. *IARC Monographs on the Evaluation of Chronic Risks to Humans*, 100(Pt C):11–465, 2009.

Illumina. Infinium HumanMethylation450 BeadChip data sheet, 2012. URL `http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_humanmethylation450.pdf`.

Illumina. Infinium MethylationEPIC BeadChip data sheet, 2015. URL `http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/humanmethylationepic-data-sheet-1070-2015-008.pdf`.

International Rice Research Institute. World rice statistics online query facility, 2011. URL `http://ricestat.irri.org:8080/wrsv3/entrypoint.htm`.

Jaenisch, R and Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics*, 33(3s):245–254, 2003. doi: 10.1038/ng1089.

Jang, H S; Shin, W J; Lee, J E, and Do, J T. CpG and non-CpG methylation in epigenetic gene regulation and brain function. *Genes*, 8(6), 2017. doi: 10.3390/genes8060148.

Jansen, R J; Argos, M; Tong, L; Li, J; Rakibuz-Zaman, M; Islam, M T; Slavkovich, V; Ahmed, A; Navas-Acien, A; Parvez, F; Chen, Y; Gamble, M V; Graziano, J H; Pierce, B L, and Ahsan, H. Determinants and Consequences of Arsenic Metabolism Efficiency among 4,794 Individuals: Demographics, Lifestyle, Genetics, and Toxicity. *Cancer Epidemiology Biomarkers and Prevention*, 25(2):381–390, 2016. doi: 10.1158/1055-9965.EPI-15-0718.

Jin, S-G; Wu, X; Li, A X, and Pfeifer, G P. Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Research*, 39(12):5015–5024, 2011. doi: 10.1093/nar/gkr120.

Kajantie, E; Osmond, C; Barker, D JP; Forsén, T; Phillips, D IW, and Eriksson, J G. Size at birth as a predictor of mortality in adulthood: a follow-up of 350,000 person-years. *International Journal of Epidemiology*, 34(3):655–663, 2005. doi: 10.1093/ije/dyi048.

Kamynina, E; Lachenauer, E R; DiRisio, A C; Liebenthal, R P; Field, M S, and Stover, P J. Arsenic trioxide targets MTHFD1 and SUMO-dependent nuclear *de novo* thymidylate biosynthesis. *PNAS*, 114(12):E2319–E2326, 2017. doi: 10.1073/pnas.1619745114.

Kaushal, A; Zhang, H; Karmaus, W J J; Everson, T M; Marsit, C J; Karagas, M R; Tsai, S-F; Wen, H-J, and Wang, S-L. Genome-wide DNA methylation at birth in relation to *in utero* arsenic exposure and the associated health in later life. *Environmental Health*, 16(1):50, 2017. doi: 10.1186/s12940-017-0262-0.

Kile, M L; Hoffman, E; Hsueh, Y-M; Afroz, S; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Ryan, L, and Christiani, D C. Variability in biomarkers of arsenic exposure and metabolism in adults over time. *Environmental Health Perspectives*, 117(3):455–460, 2009. doi: 10.1289/ehp.11251.

Kile, M L; Baccarelli, A; Hoffman, E; Tarantini, L; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Mostofa, G; Hsueh, Y-M; Wright, R O, and Christiani, D C. Prenatal arsenic exposure and DNA methylation in maternal and umbilical cord blood leukocytes. *Environmental Health Perspectives*, 120(7):1061–1066, 2012. doi: 10.1289/ehp.1104173.

Kile, M L; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Cardenas, A; Wright, R O, and Christiani, D C. Effect of prenatal arsenic exposure on DNA methylation and leukocyte subpopulations in cord blood. *Epigenetics*, 9(5):774–82, 2014. doi: 10.4161/epi.28153.

Kile, M L; Cardenas, A; Rodrigues, E; Mazumdar, M; Dobson, C; Golam, M; Quamruzzaman, Q; Rahman, M, and Christiani, D C. Estimating effects of arsenic exposure during pregnancy on perinatal outcomes in a Bangladeshi cohort. *Epidemiology*, 27 (2):1, 2015. doi: 10.1097/EDE.0000000000000416.

Kitchin, K T and Wallace, K. The role of protein binding of trivalent arsenicals in arsenic carcinogenesis and toxicity. *Journal of Inorganic Biochemistry*, 102(3):532–539, 2008. doi: 10.1016/j.jinorgbio.2007.10.021.

Klengel, T; Pape, J; Binder, E B, and Mehta, D. The role of DNA methylation in stress-related psychiatric disorders. *Neuropharmacology*, 80:115–132, 2014. doi: 10.1016/j.neuropharm.2014.01.013.

Knop, M R; Geng, TT; Gorny, A W; Ding, R; Li, C; Ley, S H, and Huang, T. Birth weight and risk of type 2 miabetes Mellitus, cardiovascular disease, and hypertension in adults: A metaanalysis of 7,646,267 participants from 135 studies. *Journal of the American Heart Association*, 7(23), 2018. doi: 10.1161/JAHA.118.008870.

Kobayashi, Y and Hirano, S. Effects of endogenous hydrogen peroxide and glutathione on the stability of arsenic metabolites in rat bile. *Toxicology and Applied Pharmacology*, 232(1):33–40, 2008. doi: 10.1016/j.taap.2008.06.003.

Koestler, D C; Avissar-Whiting, M; Houseman, E A; Karagas, M R, and Marsit, C J. Differential DNA methylation in umbilical cord blood of infants exposed to low levels of arsenic *in utero*. *Environmental Health Perspectives*, 121(8):971–7, 2013. doi: 10.1289/ehp.1205925.

Kozul-Horvath, C D; Zandbergen, F; Jackson, B P; Enelow, R I, and Hamilton, J W. Effects of low-dose drinking water arsenic on mouse fetal and postnatal growth and development. *PLoS ONE*, 7(5):e38249, 2012. doi: 10.1371/journal.pone.0038249.

Kumarathilaka, P; Seneweera, S; O, Yong S; Meharg, A, and Bundschuh, J. Arsenic in cooked rice foods: assessing health risks and mitigation options. *Environment International*, 127:584–591, 2019. doi: 10.1016/j.envint.2019.04.004.

Kuo, C-C; Moon, K A; Wang, S-L; Silbergeld, E, and Navas-Acien, A. The association of arsenic metabolism with cancer, cardiovascular disease, and diabetes: a systematic review of the epidemiological evidence. *Environmental Health Perspectives*, 125(8), 2017. doi: 10.1289/EHP577.

Kurzius-Spencer, M; da Silva, V; Thomson, C A; Hartz, V; Hsu, C-H; Burgess, J L; O'Rourke, M K, and Harris, R B. Nutrients in one-carbon metabolism and urinary arsenic methylation in the National Health and Nutrition Examination Survey (NHANES) 20032004. *Science of The Total Environment*, 607-608:381–390, 2017. doi: 10.1016/j.scitotenv.2017.07.019.

Laine, J E; Bailey, K A; Rubio-Andrade, M; Olshan, A F; Smeester, L; Drobná, Z; Herring, A H; Stýblo, M; García-Vargas, G G, and Fry, R C. Maternal arsenic exposure, arsenic methylation efficiency, and birth outcomes in the Biomarkers of Exposure to ARsenic (BEAR) Pregnancy cohort in Mexico. *Environmental Health Perspectives*, 123(2):186–192, 2015. doi: 10.1289/ehp.1307476.

Laurent, L; Wong, E; Li, G; Huynh, T; Tsirigos, A; Ong, C T; Low, H M; Kin Sung, K W; Rigoutsos, I; Loring, J, and Wei, C-L. Dynamic changes in the human methylome during differentiation. *Genome Research*, 20(3):320–31, 2010. doi: 10.1101/gr.101907.109.

Lawley, S D; Yun, J; Gamble, M V; Hall, M N; Reed, M C, and Nijhout, H. Mathematical modeling of the effects of glutathione on arsenic methylation. *Theoretical Biology and Medical Modelling*, 11(1):20, 2014. doi: 10.1186/1742-4682-11-20.

Li, A; Zhuang, T; Shi, J; Liang, Y, and Song, M. Heavy metals in maternal and cord blood in Beijing and their efficiency of placental transfer. *Journal of Environmental Sciences (China)*, 80:99–106, 2019. doi: 10.1016/j.jes.2018.11.004.

Li, L; Ekström, E-C; Goessler, W; Lönnerdal, B; Nermell, B; Yunus, M; Rahman, A; El Arifeen, S; Persson, L Å, and Vahter, M. Nutritional status has marginal influence on the metabolism of inorganic arsenic in pregnant Bangladeshi women. *Environmental Health Perspectives*, 116(3):315–321, 2007. doi: 10.1289/ehp.10639.

Li, X; Li, B; Xi, S; Zheng, Q; Lv, X, and Sun, G. Prolonged environmental exposure of arsenic through drinking water on the risk of hypertension and type 2 diabetes. *Environmental Science and Pollution Research*, 20(11):8151–8161, 2013a. doi: 10.1007/s11356-013-1768-9.

Li, X; Li, B; Xi, S; Zheng, Q; Wang, D, and Sun, G. Association of urinary monomethylated arsenic concentration and risk of hypertension: a cross-sectional study from arsenic contaminated areas in northwestern China. *Environmental Health*, 12(1):37, 2013b. doi: 10.1186/1476-069X-12-37.

Li, Y; Wang, D; Li, X; Zheng, Q, and Sun, G. A potential synergy between incomplete arsenic methylation capacity and demographic characteristics on the risk of hypertension: findings from a cross-sectional study in an arsenic-endemic area of inner Mongolia, China. *International Journal of Environmental Research and Public Health*, 12(4):3615–32, 2015. doi: 10.3390/ijerph120403615.

Liao, K-W; Chang, C-H; Tsai, M-S; Chien, L-C; Chung, M-Y; Mao, I-F; Tsai, Y-A, and Chen, M-L. Associations between urinary total arsenic levels, fetal development, and neonatal birth outcomes: a cohort study in Taiwan. *Science of the Total Environment*, 612:1373–1379, 2018. doi: 10.1016/j.scitotenv.2017.08.312.

Lin, Pi-I D; Bromage, S; Mostofa, M G; Rahman, M; Allen, J; Oken, E; Kile, M L, and Christiani, D C. Mediating role of arsenic in the relationship between diet and pregnancy outcomes: prospective birth cohort in Bangladesh. *Environmental Health*, 18(1):10, 2019. doi: 10.1186/s12940-019-0450-1.

Lin, S; Shi, Q; Nix, F B; Styblo, M; Beck, M A; Herbin-Davis, K M; Hall, L L; Simeonsson, J B, and Thomas, D J. A novel *S*-adenosyl-L-methionine:arsenic(III) methyltransferase from rat liver cytosol. *Journal of Biological Chemistry*, 277(13):10795–10803, 2002. doi: 10.1074/jbc.M110246200.

Lindberg, A-L; Rahman, M; Persson, L, and Vahter, M. The risk of arsenic induced skin lesions in Bangladeshi men and women is affected by arsenic metabolism and the age at first exposure. *Toxicology and Applied Pharmacology*, 230(1):9–16, 2008. doi: 10.1016/j.taap.2008.02.001.

Lister, R; Pelizzola, M; Dowen, R H; Hawkins, R D; Hon, G; TTonti-Filippini, J; Nery, J R; Lee, L; Ye, Z; Ngo, Q-M; Edsall, L; Antosiewicz-Bourget, J; Stewart, R; Ruotti, V; Millar, A H; Thomson, J A; Ren, B, and Ecker, J R. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462(7271):315–22, 2009. doi: 10.1038/nature08514.

Liu, H; Lu, S; Zhang, B; Xia, W; Liu, W; Peng, Y; Zhang, H; Wu, K; Xu, S, and Li, Y. Maternal arsenic exposure and birth outcomes: a birth cohort study in Wuhan, China. *Environmental Pollution*, 236:817–823, 2018. doi: 10.1016/j.envpol.2018.02.012.

Liu, S; Jiang, J; Li, L; Amato, N J; Wang, Z, and Wang, Y. Arsenite targets the zinc finger domains of Tet proteins and inhibits Tet-mediated oxidation of 5-methylcytosine. *Environmental Science and Technology*, 49(19):11923–11931, 2015. doi: 10.1021/acs.est.5b03386.

Liu, X; Zheng, Y; Zhang, W; Zhang, X; Lloyd-Jones, Dd M; Baccarelli, A A; Ning, H; Fornage, M; He, K; Liu, K, and Hou, L. Blood methylomics in response to arsenic exposure in a low-exposed US population. *Journal of Exposure Science and Environmental Epidemiology*, 24(2):145–149, 2014. doi: 10.1038/jes.2013.89.

López-Carrillo, L; Hernández-Ramírez, R U; Gandolfi, A J; Ornelas-Aguirre, J M; Torres-Sánchez, L, and Cebrian, M E. Arsenic methylation capacity is associated with breast cancer in northern Mexico. *Toxicology and Applied Pharmacology*, 280 (1):53–59, 2014. doi: 10.1016/j.taap.2014.07.013.

López-Carrillo, L; Gamboa-Loira, B; Becerra, W; Hernández-Alcaraz, C; Hernández-Ramírez, R U; Gandolfi, A J; Franco-Marina, F, and Cebrián, M E. Dietary micronutrient intake and its relationship with arsenic metabolism in Mexican women. *Environmental Research*, 151:445–450, 2016. doi: 10.1016/j.envres.2016.08.015.

Lu, H; Liu, X; Deng, Y, and Qing, H. DNA methylation, a hand behind neurodegenerative diseases. *Frontiers in Aging Neuroscience*, 5:85, 2013. doi: 10.3389/fnagi.2013.00085.

Luo, Y; McCullough, L E; Tzeng, J-Y; Darrah, T; Vengosh, A; Maguire, R L; Maity, A; Samuel-Hodge, C; Murphy, S K; Mendez, M A, and Hoyo, C. Maternal blood cadmium, lead and arsenic levels, nutrient combinations, and offspring birthweight. *BMC Public Health*, 17(1):354, 2017. doi: 10.1186/s12889-017-4225-8.

Mandal, B K; Ogra, Y, and Suzuki, K T. Identification of dimethylarsinous and monomethylarsonous acids in human urine of the arsenic-affected areas in West Bengal, India. *Chemical Research in Toxicology*, 14(4):371–8, 2001. doi: 10.1021/TX000246H.

Marshall, G; Ferreccio, C; Yuan, Y; Bates, M N; Steinmaus, C; Selvin, S; Liaw, J, and Smith, A H. Fifty-year study of lung and bladder cancer mortality in Chile related to arsenic in drinking water. *Journal of the National Cancer Institute*, 99(12):920–8, 2007. doi: 10.1093/jnci/djm004.

Mazumdar, M; Ibne Hasan, M O S; Hamid, R; Valeri, L; Paul, L; Selhub, J; Rodrigues, E G; Silva, F; Mia, S; Mostofa, M G; Quamruzzaman, Q; Rahman, M, and Christiani, D C. Arsenic is associated with reduced effect of folic acid in myelomeningocele prevention: a case control study in Bangladesh. *Environmental Health*, 14(1):34, 2015. doi: 10.1186/s12940-015-0020-0.

Meharg, A A; Lombi, E; Williams, P N; Scheckel, K G; Feldmann, J; Raab, A; Zhu, Y, and Islam, R. Speciation and localization of arsenic in white and brown rice grains. *Environmental Science ane Technology*, 42(4):1051–7, 2008.

Melak, D; Ferreccio, C; Kalman, D; Parra, R; Acevedo, J; Pérez, L; Cortés, S; Smith, A H; Yuan, Y; Liaw, J, and Steinmaus, C. Arsenic methylation and lung and bladder cancer in a case-control study in northern Chile. *Toxicology and Applied Pharmacology*, 274(2):225–31, 2014. doi: 10.1016/j.taap.2013.11.014.

Mellén, M; Ayata, P; Dewell, S; Kriaucionis, S, and Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell*, 151(7):1417–30, 2012. doi: 10.1016/j.cell.2012.11.022.

Milton, A H; Hussain, S; Akter, S; Rahman, M; Mouly, T A, and Mitchell, K. A review of the effects of chronic arsenic exposure on adverse pregnancy outcomes. *International Journal of Environmental Research and Public Health*, 14(6), 2017. doi: 10.3390/ijerph14060556.

Mitra, S R; Mazumder, D N G; Basu, A; Block, G; Haque, R; Samanta, S; Ghosh, N; Smith, M M H; von Ehrenstein, O S, and Smith, A H. Nutritional factors and susceptibility to arsenic-caused skin lesions in West Bengal, India. *Environmental Health Perspectives*, 112(10):1104–9, 2004. doi: 10.1289/EHP.6841.

Moe, B; Peng, H; Lu, X; Chen, B; Chen, L WL; Gabos, S; Li, X-F, and Le, X C. Comparative cytotoxicity of fourteen trivalent and pentavalent arsenic species determined using real-time cell sensing. *Journal of Environmental Sciences*, 49:113–124, 2016. doi: 10.1016/j.jes.2016.10.004.

Moon, K A; Oberoi, S; Barchowsky, A; Chen, Y; Guallar, E; Nachman, K E; Rahman, M; Sohel, N; D'Ippoliti, D; Wade, T J; James, K A; Farzan, S F; Karagas, M R; Ahsan, H, and Navas-Acien, A. A dose-response meta-analysis of chronic arsenic exposure and incident cardiovascular disease. *International Journal of Epidemiology*, 46(6):1924–1939, 2017. doi: 10.1093/ije/dyx202.

Moore, C L; Flanigan, T J; Law, C D; Loukotková, L; Woodling, K A; da Costa, G G; Fitzpatrick, S C, and Ferguson, S A. Developmental neurotoxicity of inorganic arsenic exposure in Sprague-Dawley rats. *Neurotoxicology and Teratology*, 72:49–57, 2019. doi: 10.1016/j.ntt.2019.01.007.

Moran, S; Arribas, C, and Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics*, 6:epi.15.114, 2015. doi: 10.2217/epi.15.114.

Morrissey, R E and Mottet, N K. Arsenic-induced exencephaly in the mouse and associated lesions occurring during neurulation. *Teratology*, 28(3):399–411, 1983. doi: 10.1002/tera.1420280311.

Mudd, S H and Poole, J R. Labile methyl balances for normal humans on various dietary regimens. *Metabolism*, 24(6):721–735, 1975. doi: 10.1016/0026-0495(75)90040-2.

Myers, S L; Lobdell, D T; Liu, Z; Xia, Y; Ren, H; Li, Y; Kwok, R K; Mumford, J L, and Mendola, P. Maternal drinking water arsenic exposure and perinatal outcomes in inner Mongolia, China. *Journal of Epidemiology and Community Health*, 64(4): 325–329, 2010. doi: 10.1136/jech.2008.084392.

Nachman, K E; Baron, P A; Raber, G; Francesconi, K A; Navas-Acien, A, and Love, D C. Roxarsone, inorganic arsenic, and other arsenic species in chicken: a U.S.-based market basket sample. *Environmental Health Perspectives*, 121(7):818–24, 2013. doi: 10.1289/ehp.1206245.

Nagymajtényi, L; Selypes, A, and Berencsi, G. Chromosomal aberrations and fetotoxic effects of atmospheric arsenic exposure in mice. *Journal of Applied Toxicology*, 5(2): 61–3, 1985.

National Research Council. Distribution of arsenic in the environment. In *Arsenic: medical and biologic effects of environmental pollutants.*, pages 16–79. National Academies Press, Washington, DC, 1977.

National Research Council. Critical aspects of EPA's and IRIS assessment of inorganic arsenic, interim report. Technical report, National Academies, Washington, DC, 2013.

Navas-Acien, A; Umans, J G; Howard, B V; Goessler, Q; Francesconi, K A; Crainiceanu, C M; Silbergeld, E K, and Guallar, E. Urine arsenic concentrations and species excretion patterns in American Indian communities over a 10-year period: The Strong Heart Study. *Environmental Health Perspectives*, 117(9):1428–1433, 2009. doi: 10.1289/ehp.0800509.

Neidhart, M and Neidhart, M. DNA methylation and autoimmunity. In *DNA Methylation and Complex Human Disease*, chapter 19, pages 327–340. 2016. ISBN 9780124201941. doi: 10.1016/B978-0-12-420194-1.00019-1.

Niedzwiecki, M M; Hall, M N; Liu, X; Oka, J; Harper, K N; Slavkovich, V; Ilievski, V; Levy, D; van Geen, A; Mey, J L; Alam, S; Siddique, A B; Parvez, F; Graziano, J H, and Gamble, M V. A dose-response study of arsenic exposure and global methylation of peripheral blood mononuclear cell DNA in Bangladeshi adults. *Environmental Health Perspectives*, 121(11-12):1306–12, 2013. doi: 10.1289/ehp.1206421.

Niedzwiecki, M M; Liu, X; Hall, M N; Thomas, T; Slavkovich, V; Ilievski, V; Levy, D; Alam, S; Siddique, A B; Parvez, F; Graziano, J H, and Gamble, M V. Sex-specific associations of arsenic exposure with global DNA methylation and hydroxymethylation in leukocytes: results from two studies in Bangladesh. *Cancer Epidemiology, Biomarkers and Prevention*, 24(11):1748–57, 2015. doi: 10.1158/1055-9965.EPI-15-0432.

Niedzwiecki, M M; Liu, X; Zhu, H; Hall, M N; Slavkovich, V; Ilievski, V; Levy, D; Siddique, A B; Kibriya, M G; Parvez, F; Islam, T; Ahmed, A; Navas-Acien, A;

Graziano, J H; Finnell, R H; Ahsan, H, and Gamble, M V. Serum homocysteine, arsenic methylation, and arsenic-induced skin lesion incidence in Bangladesh: a one-carbon metabolism candidate gene study. *Environment International*, 113:133–142, 2018. doi: 10.1016/j.envint.2018.01.015.

Nilsson, E; Jansson, P A; Perfilyev, A; Volkov, P; Pedersen, M; Svensson, M K; Poulsen, P; Ribel-Madsen, R; Pedersen, N L; Almgren, P; Fadista, J; Rönn, T; Klarlund Pedersen, B; Scheele, C; Vaag, A, and Ling, C. Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. *Diabetes*, 63(9):2962–76, 2014. doi: 10.2337/db13-1459.

Nriagu, J O; Bhattacharya, P; Mukherjee, A B; Bundschuh, J; Zevenhoven, R, and Loeppert, R H. Arsenic in soil and groundwater: an overview. *Trace Metals and other Contaminants in the Environment*, 9:3–60, 2007. doi: 10.1016/S1875-1121(06)09001-8.

Okano, M; Bell, D W; Haber, D A, and Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell*, 99(3):247–57, 1999.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Sanchez, T; van Geen, A; Mey, J L; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T; Balac, O; Ilievski, V; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folic acid and creatine as therapeutic approaches to lower blood arsenic: a randomized controlled trial. *Environmental Health Perspectives*, 123(12):1294–1301, 2015a. doi: 10.1289/ehp.1409396.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M Nasir; Islam, T; Ilievski, V; Graziano, J H, and Gamble, M V. Low-dose creatine supplementation lowers plasma guanidinoacetate, but not plasma homocysteine, in a double-blind, randomized, placebo-controlled trial. *Journal of Nutrition*, 145(10):2245–52, 2015b. doi: 10.3945/jn.115.216739.

Petrick, J S; Ayala-Fierro, F; Cullen, W R; Carter, D E, and Vasken Aposhian, H. Monomethylarsonous acid (MMA(III)) is more toxic than arsenite in Chang human hepatocytes. *Toxicology and Applied Pharmacology*, 163(2):203–207, 2000. doi: 10.1006/taap.1999.8872.

Petrick, J S; Jagadish, B; Mash, E A, and Aposhian, H V. Monomethylarsonous acid (MMA^{III}) and arsenite: LD50 in hamsters and *in vitro* inhibition of pyruvate dehydrogenase. *Chemical Research in Toxicology*, 14(6):651–56, 2001. doi: 10.1021/TX000264Z.

Pilsner, J R; Liu, X; Ahsan, H; Ilievski, V; Slavkovich, V; Levy, D; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Genomic methylation of peripheral blood leukocyte DNA: influences of arsenic and folate in Bangladeshi adults. *American Journal of Clinical Nutrition*, 86(4):1179–86, 2007.

Pilsner, J R; Liu, X; Ahsan, H; Ilievski, V; Slavkovich, V; Levy, D; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. *Environmental Health Perspectives*, 117(2):254–60, 2009. doi: 10.1289/ehp.11872.

Potera, C. U.S. rice serves up arsenic. *Environmental Health Perspectives*, 115(6):A296, 2007. doi: 10.1289/ehp.115-a296.

Pu, Y-S; Yang, S-M; Huang, Y-K; Chung, C-J; Huang, S K; Chiu, A W-H; Yang, M-H; Chen, C-J, and Hsueh, Y-M. Urinary arsenic profile affects the risk of urothelial carcinoma even at low arsenic exposure. *Toxicology and Applied Pharmacology*, 218 (2):99–106, 2007. doi: 10.1016/j.taap.2006.09.021.

Putiri, E L and Robertson, K D. Epigenetic mechanisms and genome stability. *Clinical Epigenetics*, 2(2):299–314, 2011. doi: 10.1007/s13148-010-0017-z.

Rahman, A; Vahter, M; Smith, A H; Nermell, B; Yunus, M; El Arifeen, S; Persson, L-A, and Ekström, E-C. Arsenic exposure during pregnancy and size at birth: a prospective cohort study in Bangladesh. *American Journal of Epidemiology*, 169(3): 304–312, 2008. doi: 10.1093/aje/kwn332.

Rahman, M L; Valeri, L; Kile, M L; Mazumdar, M; Mostofa, G; Qamruzzaman, Q; Rahman, M; Baccarelli, A; Liang, L; Hauser, R, and Christiani, D C. Investigating causal relation between prenatal arsenic exposure and birthweight: are smaller infants more susceptible? *Environment International*, 108:32–40, 2017.

Raml, R; Rumpler, A; Goessler, W; Vahter, M; Li, L; Ochi, T, and Francesconi, K A. Thio-dimethylarsinate is a common metabolite in urine samples from arsenic-exposed women in Bangladesh. *Toxicology and Applied Pharmacology*, 222(3):374–380, 2007. doi: 10.1016/j.taap.2006.12.014.

Ravenscroft, P; Brammer, H, and Richards, K. *Arsenic pollution: a global synthesis*. Wiley-Blackwell, Hoboken, NJ, 2009. ISBN 9781444308785. doi: 10.1002/9781444308785.

Reichard, J F and Puga, A. Effects of arsenic exposure on DNA methylation and epigenetic gene regulation. *Epigenomics*, 2(1):87–104, 2010. doi: 10.2217/epi.09.45.

Remy, S; Govarts, E; Bruckers, L; Paulussen, M; Wens, B; Hond, E Den; Nelen, V; Baeyens, W; van Larebeke, N; Loots, I; Sioen, I, and Schoeters, G. Expression of the sFLT1 gene in cord blood cells is associated to maternal arsenic exposure and decreased birth weight. *PloS ONE*, 9(3):e92677, 2014. doi: 10.1371/journal.pone.0092677.

Robertson, K D. DNA methylation and human disease. *Nature Reviews. Genetics*, 6 (August):597–610, 2005. doi: 10.1038/nrg1655.

Rojas, D; Rager, J E; Smeester, L; Bailey, K A; Drobná, Z; Rubio-Andrade, M; Stýblo, M; García-Vargas, G, and Fry, R C. Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. *Toxicological Sciences*, 143(1):97–106, 2015. doi: 10.1093/toxsci/kfu210.

Röllin, H B; Channa, K; Olutola, B G, and Odland, J Ø. Evaluation of *in utero* exposure to arsenic in South Africa. *Science of the Total Environment*, 575:338–346, 2017. doi: 10.1016/j.scitotenv.2016.10.044.

Saxonov, S; Berg, P, and Brutlag, D L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *PNAs*, 103(5): 1412–1417, 2006. doi: 10.1073/pnas.0510310103.

Shankar, S; Shanker, U, and Shikha. Arsenic contamination of groundwater: a review of sources, prevalence, health risks, and strategies for mitigation. *Scientific World Journal*, 2014:304524, 2014. doi: 10.1155/2014/304524.

Smith, A H; Lingas, E O, and Rahman, M. Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. *Bull World Health Organ*, 78(9):1093–103, 2000.

Smith, A H; Marshall, G; Liaw, J; Yuan, Y; Ferreccio, C, and Steinmaus, C. Mortality in young adults following *in utero* and childhood exposure to arsenic in drinking water. *Environmental Health Perspectives*, 120(11):1527–31, 2012. doi: 10.1289/ehp.1104867.

Smith, Z D and Meissner, A. DNA methylation: roles in mammalian development. *Nature Reviews Genetics*, 14(3):204–220, 2013. doi: 10.1038/nrg3354.

Spiegelstein, O; Lu, X; Le, X C; Troen, A; Selhub, J; Melnyk, S; James, S J, and Finnell, R H. Effects of dietary folate intake and folate binding protein-1 (Folbp1) on urinary speciation of sodium arsenate in mice. *Toxicology Letters*, 145(2):167–174, 2003. doi: 10.1016/S0378-4274(03)00307-2.

Spiegelstein, O; Gould, A; Wlodarczyk, B; Tsie, M; Lu, X; Le, C; Troen, A; Selhub, J; Piedrahita, J A; Salbaum, J M; Kappen, C; Melnyk, S; James, J, and Finnell, R H. Developmental consequences of *in utero* sodium arsenate exposure in mice with folate transport deficiencies. *Toxicology and Applied Pharmacology*, 203(1):18–26, 2005. doi: 10.1016/j.taap.2004.07.006.

Spratlen, M J; Gamble, M V; Grau-Perez, M; Kuo, C-C; Best, L G; Yracheta, J; Francesconi, K; Goessler, W; Mossavar-Rahmani, Y; Hall, M; Umans, J G; Fretts, A, and Navas-Acien, A. Arsenic metabolism and one-carbon metabolism at low-moderate arsenic exposure: evidence from the Strong Heart Study. *Food and Chemical Toxicology*, 105:387–397, 2017. doi: 10.1016/j.fct.2017.05.004.

Stead, L M; Au, K P; Jacobs, R L; Brosnan, M E, and Brosnan, J T. Methylation demand and homocysteine metabolism: effects of dietary provision of creatine and guanidinoacetate. *American Journal of Physiology*, 281(5):E1095–100, 2001. doi: 10.1152/ajpendo.2001.281.5.E1095.

Stead, L M; Brosnan, J T; Brosnan, M E; Vance, D E, and Jacobs, R. Is it time to reevaluate methyl balance in humans? *American Journal of Clinical Nutrition*, 83 (1):5–10, 2006.

Steinmaus, C; Carrigan, K; Kalman, D; Atallah, R; Yuan, Y, and Smith, A H. Dietary intake and arsenic methylation in a US population. *Environmental Health Perspectives*, 113(9):1153–9, 2005.

Steinmaus, C; Bates, M N.; Yuan, Y; Kalman, D; Atallah, R; Rey, O A; Biggs, M L; Hopenhayn, C; Moore, L E; Hoang, B K, and Smith, A H. Arsenic methylation and bladder cancer risk in case-control studies in Argentina and the United States. *Journal of Occupational and Environmental Medicine*, 48(5):478–488, 2006. doi: 10.1097/01.jom.0000200982.28276.70.

Steinmaus, C; Yuan, Y; Kalman, D; Rey, O A; Skibola, C F; Dauphine, D; Basu, A; Porter, K E; Hubbard, A; Bates, M N; Smith, M T, and Smith, A H. Individual differences in arsenic metabolism and lung cancer in a case-control study in Cordoba, Argentina. *Toxicology and Applied Pharmacology*, 247(2):138–45, 2010. doi: 10.1016/j.taap.2010.06.006.

Steinmaus, C; Ferreccio, C; Acevedo, J; Yuan, Y; Liaw, J; Duran, V; Cuevas, S; Garcia, J; Meza, R; Valdes, R; Valdes, G; Benitez, H; VanderLinde, V; Villagra, V; Cantor, K P; Moore, L E; Perez, S G; Steinmaus, S, and Smith, A H. Increased lung and bladder cancer incidence in adults after *in utero* and early-life arsenic exposure. *Cancer Epidemiology Biomarkers and Prevention*, 23(8):1529–38, 2014. doi: 10.1158/1055-9965.EPI-14-0059.

Steinmaus, C M; Ferreccio, C; Romo, J A; Yuan, Y; Cortes, S; Marshall, G; Moore, L E; Balmes, J R; Liaw, J; Golden, T, and Smith, A H. Drinking water arsenic in northern Chile: high cancer risks 40 years after exposure cessation. *Cancer Epidemiology Biomarkers and Prevention*, 22(4):623–30, 2013. doi: 10.1158/1055-9965.EPI-12-1190.

Stroud, H; Feng, S; Morey Kinney, S; Pradhan, S, and Jacobsen, S E. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biology*, 12(6):R54, 2011. doi: 10.1186/gb-2011-12-6-r54.

Styblo, M; Del Razo, L M; Vega, L; Germolec, D R; LeCluyse, E L; Hamilton, G A; Reed, W; Wang, C; Cullen, W R, and Thomas, D J. Comparative toxicity of trivalent and pentavalent inorganic and methylated arsenicals in rat and human cells. *Archives of Toxicology*, 74(6):289–299, 2000. doi: 10.1007/s002040000134.

Sun, X; Liu, W; Zhang, B; Shen, X; Hu, C; Chen, X; Jin, S; Jiang, Y; Liu, H; Cao, Z; Xia, W; Xu, S, and Li, Y. Maternal heavy metal exposure, thyroid hormones, and birth outcomes: a prospective cohort study. *Journal of Clinical Endocrinology and Metabolism*, page Epub ahead of print, 2019. doi: 10.1210/jc.2018-02492.

Taes, Y E C; D, Joris R; De Vriese, A S; Rombaut, R; Van Camp, J, and Lameire, N H. Creatine supplementation decreases homocysteine in an animal model of uremia. *Kidney International*, 64(4):1331–7, 2003. doi: 10.1046/j.1523-1755.2003.00206.x.

Tajuddin, S M; Amaral, A F; Fernández, A F; Rodríguez-Rodero, S; Rodríguez, R M; Moore, L E; Tardón, A; Carrato, A; García-Closas, M; Silverman, D T; Jackson, B P; García-Closas, R; Cook, A L; Cantor, K P; Chanock, S; Kogevinas, M; Rothman, N; Real, F X; Fraga, M F; Malats, N N, and Spanish Bladder Cancer/EPICURO Study Investigators. Genetic and non-genetic predictors of LINE-1 methylation in leukocyte DNA. *Environmental Health Perspectives*, 121(6):650–656, 2013. doi: 10.1289/ehp.1206068.

Thomas, D; Waters, Stephen B, and Styblo, Miroslav. Elucidating the pathway for arsenic methylation. *Toxicology and Applied Pharmacology*, 198(3):319–326, 2004. doi: 10.1016/j.taap.2003.10.020.

Tice, R R; Yager, J W; Andrews, P, and Crecelius, E. Effect of hepatic methyl donor status on urinary excretion and DNA damage in B6C3F1 mice treated with sodium arsenite. *Mutation Research*, 386(3):315–34, 1997. doi: 10.1016/S1383-5742(97)00004-5.

Tseng, C-H; Huang, Y-K Y-L; Chung, C-J; Yang, M-H; Chen, C-J, and Hsueh, Y-M. Arsenic exposure, urinary arsenic speciation, and peripheral vascular disease in blackfoot disease-hyperendemic villages in Taiwan. *Toxicology and Applied Pharmacology*, 206(3):299–308, 2005. doi: 10.1016/j.taap.2004.11.022.

Tyler, C R and Allan, A M. The effects of arsenic exposure on neurological and cognitive dysfunction in human and rodent studies: a review. *Current Environmental Health Reports*, 1(2):132–147, 2014. doi: 10.1007/s40572-014-0012-1.

US Environmental Protection Agency. Drinking water arsenic rule history, 2016.

US EPA Office of Water. Drinking water standard for arsenic. Technical report, US Environmental Protection Agency, 2001. URL `https://nepis.epa.gov/Exe/ZyPdf.cgi?Dockey=20001XXC.txt`.

Vahter, M. Effects of arsenic on maternal and fetal health. *Annual Review of Nutrition*, 29(1):381–399, 2009. doi: 10.1146/annurev-nutr-080508-141102.

Vahter, M and Marafante, E. Effects of low dietary intake of methionine, choline or proteins on the biotransformation of arsenite in the rabbit. *Toxicology Letters*, 37(1):41–46, 1987. doi: 10.1016/0378-4274(87)90165-2.

Valenzuela, O L; Borja-Aburto, V H; Garcia-Vargas, G G; Cruz-Gonzalez, M B; Garcia-Montalvo, E A; Calderon-Aranda, E S, and Del Razo, L M. Urinary trivalent methylated arsenic species in a population chronically exposed to inorganic arsenic. *Environmental Health Perspectives*, 113(3):250–254, 2004. doi: 10.1289/ehp.7519.

Vance, J E and Vance, D E. Phospholipid biosynthesis in mammalian cells. *Biochemistry and Cell Biology*, 82(1):113–128, 2004. doi: 10.1139/o03-073.

Vijayaraghavan, M; Wanibuchi, H; Karim, R; Yamamoto, S; Masuda, C; Nakae, D; Konishi, Y, and Fukushima, S. Dimethylarsinic acid induces 8-hydroxy-2-deoxyguanosine formation in the kidney of NCI-Black-Reiter rats. *Cancer Letters*, 165(1):11–17, 2001. doi: 10.1016/S0304-3835(00)00711-4.

Wardlaw, T; Blanc, A; Zupan, J; Ahman, E, and Åhman, E. Low birthweight: country, regional, and global estimates. Technical report, New York, NY, 2004.

Wasserman, G A; Liu, X; Parvez, F; Factor-Litvak, P; Ahsan, H; Levy, D; Kline, J; van Geen, A; Mey, J; Slavkovich, V; Siddique, A B; Islam, T, and Graziano, J H. Arsenic and manganese exposure and children's intellectual function. *Neurotoxicology*, 32(4): 450–7, 2011. doi: 10.1016/j.neuro.2011.03.009.

Wasserman, G A.; Liu, X; Parvez, F; Factor-Litvak, P; Kline, J; Siddique, A B; Shahriar, H; Uddin, M N; van Geen, A; Mey, J L; Balac, O, and Graziano, J H. Child intelligence and reductions in water arsenic and manganese: a two-year follow-up study in Bangladesh. *Environmental Health Perspectives*, 124(7):1114–1120, 2016. doi: 10.1289/ehp.1509974.

Wei, Y; Shi, Q; Wang, Z; Zhang, R; Su, L; Quamruzzaman, Q; Rahman, M; Chen, F, and Christiani, D C. Maternal/fetal metabolomes appear to mediate the impact of arsenic exposure on birth weight: a pilot study. *Journal of Exposure Science and Environmental Epidemiology*, 27(3):313–319, 2017. doi: 10.1038/jes.2016.74.

Wilson, A S; Power, B E, and Molloy, P L. DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta*, 1775(1):138–162, 2007. doi: 10.1016/j.bbcan.2006.08.007.

Wlodarczyk, B; Spiegelstein, O; Gelineau-van Waes, J; Vorce, R L; Lu, X; Le, C X, and Finnell, R H. Arsenic-induced congenital malformations in genetically susceptible folate binding protein-2 knockout mice. *Toxicology and Applied Pharmacology*, 177 (3):238–246, 2001. doi: 10.1006/taap.2001.9303.

World Health Organization. Arsenic in drinking-water: background document for the development of WHO guidelines for drinking-water quality. Technical report, Geneva, Switzerland, 2011.

World Water Assessment Programme. Water for a sustainable world: facts and figures. Technical report, United Nations, Perugia, Italy, 2015. URL `http://ggmn.e-id.nl/ggmn/GlobalOverview.html`.

Wu, M-M; Chiou, H-Y; Hsueh, Y-M; Hong, C-T; Su, C-L; Chang, S-F; Huang, W-L; Wang, H-T; Wang, Y-H; Hsieh, Y-C, and Chen, C-J. Effect of plasma homocysteine level and urinary monomethylarsonic acid on the risk of arsenic-associated carotid atherosclerosis. *Toxicology and Applied Pharmacology*, 216:168–175, 2006. doi: 10.1016/j.taap.2006.05.005.

Xu, L; Yokoyama, K; Tian, Y; Piao, F-Y; Kitamura, F; Kida, H, and Wang, P. Decrease in birth weight and gestational age by arsenic among the newborn in Shanghai, China. *[Nihon koshu eisei zasshi] Japanese Journal of Public Health*, 58(2):89–95, 2011.

Yamanaka, K; Takabayashi, F; Mizoi, M; An, Y; Hasegawa, A, and Okada, S. Oral exposure of dimethylarsinic acid, a main metabolite of inorganic arsenics, in mice leads to an increase in 8-oxo-2-deoxyguanosine level, specifically in the target organs for arsenic carcinogenesis. *Biochemical and Biophysical Research Communications*, 287(1):66–70, 2001. doi: 10.1006/bbrc.2001.5551.

Yang, C-Y; Chang, C-C; Tsai, S-S; Chuang, H-Y; Ho, C-K, and Wu, T-N. Arsenic in drinking water and adverse pregnancy outcome in an arseniasis-endemic area in northeastern Taiwan. *Environmental Research*, 91(1):29–34, 2003. doi: 10.1016/S0013-9351(02)00015-4.

Yildirim, O; Li, R; Hung, J-H; Chen, Poshen B; Dong, X; Ee, L-S; Weng, Z; Rando, O J, and Fazzio, T G. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell*, 147(7):1498–1510, 2011. doi: 10.1016/j.cell.2011.11.054.

Yu, R C; Hsu, K-H; Chen, C-J, and Froines, J R. Arsenic methylation capacity and skin cancer. *Cancer Epidemiology, Biomarkers and Prevention*, 9:1259–1262, 2000.

Zakharyan, R; Wu, Y; Bogdan, G M, and Aposhian, H V. Enzymatic methylation of arsenic compounds: assay, partial purification, and properties of arsenite methyltransferase and monomethylarsonic acid methyltransferase of rabbit liver. *Chemical Research in Toxicology*, 8(8):1029–38, 1995. doi: 10.1021/TX00050A006.

Zakharyan, R A; Ayala-Fierro, F; Cullen, W R; Carter, D M, and Aposhian, H V. Enzymatic methylation of arsenic compounds. VII. Monomethylarsonous acid (MMAIII) is the substrate for MMA methyltransferase of rabbit liver and human hepatocytes. *Toxicology and Applied Pharmacology*, 158(1):9–15, 1999. doi: 10.1006/taap.1999.8687.

Zeisel, S H. Importance of methyl donors during reproduction. *American Journal of Clinical Nutrition*, 89(2), 2009. doi: 10.3945/ajcn.2008.26811D.

Zeisel, S H and Warrier, M. Trimethylamine N-oxide, the microbiome, and heart and kidney disease. *Annual Review of Nutrition*, 37(1):157–181, 2017. doi: 10.1146/annurev-nutr-071816-064732.

Zhong, Q; Cui, Y; Wu, H; Niu, Q; Lu, X; Wang, L, and Huang, F. Association of maternal arsenic exposure with birth size: a systematic review and meta-analysis. *Environmental Toxicology and Pharmacology*, 69:129–136, 2019. doi: 10.1016/j.etap.2019.04.007.

2.10 Appendix

Table 2.1: Summary of selected epidemiological studies of the association between *in utero* As exposure and birth weight.

Study	**Location**	**Study design**	**N**	**Measure of exposure**	**Mean exposure**	**Adjusted**	**Effect size**
Sun et al. (2019)	China	PC	675	Maternal urine	GM = 20.03 μg/L	Yes	β = -28.29, P = 0.15
Li et al. (2019)	China	CS	156	Maternal blood; cord blood	Mean = 2.90 in maternal blood; 2.0 in cord blood	No	$r_{Spearman}$ = 0.107, P = 0.187 for maternal blood; $r_{Spearman}$ = 0.06 0, P = 0.467 for cord blood
Lin et al. (2019)	Bangladesh	PC	1,057	Maternal toenail	Median = 1.21 μg/g	Yes	β = -5.6, P = 0.618
Goodrich et al. (2019)	US	PC	56	Maternal urine	GM = 4.30 μg/L	Yes	β = 9.03, P = 0.86
Freire et al. (2019)	Spain	CS	327	Placenta	27.5% > LOD; Median < 0.004 ng/g	Yes	β = -38.83, 95% CI: -136.3, 58.59)
Liu et al. (2018)	China	PC	1,390	Maternal urine	Median = 21.86 μg/L for 1st trimester; 20.21 μg/L for 2nd trimester; 21.08 μg/L for 3rd trimester	Yes	β = -3.89, 95% CI: -25.21, 17.42 for 1st trimester; β = 6.79, 95% CI: -16.56, 30.13 for 2nd trimester; β = -24.27, 95% CI: -46.99, -1.55 for 3rd trimester; overall P = 0.15
Liao et al. (2018)	Taiwan	PC	130	Maternal urine	GM = 41.8 μg/L for 1st trimester; 40.0 μg/L for 2nd trimester; 40.6 μg/L for 3rd trimester	Yes	β = -123.88, P = 0.070 for 1st trimester; β = -173.26, P = 0.005 for 2nd trimester; β = -14.51, P = 0.845 for 3rd trimester

Study	Location	Study design	N	Measure of exposure	Mean exposure	Adjusted	Effect size
Rahman et al. (2017)[a]	Bangladesh	PC	1182	Drinking water, maternal toenails	Median = 2.2 μg/L for water; 1.2 μg/g for toenails	Yes	Effect via pathways independent of gestational age significant for smaller infants: β = -28.0, 95% CI: -43.8, -9.9 for 10th percentile; β = -14.9, 95% CI: -30.3, -1.7 for 20th percentile
Luo et al. (2017)	US	PC	275	Maternal blood	Median = 0.047 μg/dL for < 37 weeks gestational age; 0.043 μg/dL ≥ 37 weeks gestational age	Yes	β = 366.5, P = 0.038 for moderate (2nd tercile) vs. low exposure (1st tercile); β = 72.82, P = 0.603 for high (3rd tercile) vs. low exposure
Wei et al. (2017)[a]	Bangladesh	PC	35 infants, 20 mothers	Cord blood		Yes	β = -9.40, P = 0.783
Govarts et al. (2016)	Belgium	CS	242	Cord blood	GM = 0.561 μg/L	Yes	β = -90.61, P =0.016
Gilbert-Diamond et al. (2016)	US	PC	706	Maternal urine	Medin = 3.4 μg/L	Yes	β = -1.3, 95% CI: -35.8, 33.2
Kile et al. (2015)[a]	Bangladesh	PC	1,140	Drinking water; maternal toenail	Median = 2.3 μg/L for water; 1.46 μg/g for maternal toenails	Yes	No significant direct effect of As on birth weight

Study	Location	Study design	N	Measure of exposure	Mean exposure	Adjusted	Effect size
Bloom et al. (2016)	Romania	PC	122	Drinking water	Mean (SD): 4.11 (12.99) μg/L	Yes	β = -0.06, 95% CI: -0.21, 0.08
Hu et al. (2015)	China	CS	81	Maternal blood; cord blood	Median =11.0 ng/g for maternal blood; 10.4 ng/g for cord blood	Yes	β = -1.5, P = 0.878 for maternal blood; β = -13.6, P = 0.185 for cord blood
Laine et al. (2015)	Mexico	PC	200	Maternal urine	Median = 23.3 μg/L	Yes	β = -0.58, P = 0.44
Remy et al. (2014)	Belgium	CS	177	Maternal blood; cord blood	GM = 0.56 μg/L for cord blood	Yes	β = -47, P = 0.0033 for an IQR increase of cord bAs
Guan et al. (2012)	China	CS	125	Maternal blood; cord blood	Mean = 6.91 μg/L in maternal blood; 5.41 μg/L in cord blood	Yes	β = -0.02, P = 0.015
Xu et al. (2011)	China	Cross sectional	142	Maternal blood; cord blood	Mean = 4.13 μg/L in maternal blood; 1.82 μg/L in cord blood	Yes	β = -354.41, 95% CI: -677.53, -31.28 for maternal blood in males; not significant in females
Myers et al. (2010)	China	CS	9,890	Drinking water from maternal subvillage	14% drinking water As > 50 mg/L	Yes	Mean difference for > 100 μg/L vs. > 20 μg/L As = 0.05, 95% CI: 0.02, 0.08
Rahman et al. (2008)	Bangladesh	PC	1,578	Maternal urine	Mean (SD) = 160 (163) μg/L	Yes	β = -1.68, P = 0.007 for maternal urinary As < 100 μg/L; not significant in full exposure range

Study	Location	Study design	N	Measure of exposure	Mean exposure	Adjusted	Effect size
Huyck et al. (2007)[a]	Bangladesh	PC	43	Maternal hair	Range = 0.14, 3.28 μg/g	Yes	β = -193.5, P = 0.04
Hopenhayn et al. (2003)	Chile	PC	844	City drinking water	Mean = 42 μg/L in exposed city; <1 μg/L in unexposed city	Yes	Mean difference for exposed vs. unexposed = -57, 95% CI: -123, 9
Yang et al. (2003)	Taiwan	CS	18,259	Drinking water	Range in exposed areas: < 0.15 ppb to 3.59 ppm	Yes	Mean difference for exposed vs. unexposed = -29.05, 95% CI: -44.55, -13.55

PC, Prospective cohort; CS, Cross-sectional; GM = geometric mean. a. Analyses of the same birth cohort recruited in the Sirajdikhan and Pabna Sadar upazilas of Bangladesh.

Table 2.2: Summary of selected epidemiological studies of the association between *in utero* As exposure and gestational age.

Study	**Location**	**Study design**	**N**	**Measure of exposure**	**Mean exposure**	**Adjusted**	**Effect size**
Sun et al. (2019)	China	PC	675	Maternal urine	GM = 20.03 μg/L	Yes	$\beta = 0.00$, $P = 0.94$
Lin et al. (2019)[a]	Bangladesh	PC	1,057	Maternal toenails	Median = 1.21 μg/g	Yes	$\beta = -0.17$, $P = 0.001$
Freire et al. (2019)	Spain	CS	327	Placenta	27.5% >LOD; Median < 0.004 ng/g	Yes	$\beta = 0.17$, 95% CI: -0.20, 0.54
Bloom et al. (2016)	Romania	PC	122	Drinking water	Mean: 4.11 μg/L	Yes	$\beta = 0.002$, 95% CI: -0.01, 0.02
Röllin et al. (2017)	South Africa	PC	650	Maternal blood; cord blood	GM = 0.62 μg/L for maternal blood	Yes	$\beta = -0.054$, $P = 0.002$
Kile et al. (2015)[a]	Bangladesh	PC	1,140	Drinking water; maternal toenail	Median = 2.3 μg/L for water; 1.46 μg/g for maternal toenails	Yes	Direct effect = -0.23, 95% CI: -0.28, -0.17
Laine et al. (2015)	Mexico	PC	200	Maternal urine	Median = 23.3 μg/L	Yes	$\beta = -0.0036$, $P = 0.07$
Xu et al. (2011)	China	CS	142	Maternal blood; cord blood	Mean = 4.13 μg/L in maternal blood; 1.82 μg/L in cord blood	Yes	$\beta = -1.51$, 95% CI: -2.50, -0.51 for maternal blood in males; association not significant in females

PC, Prospective cohort; CS, Cross-sectional; GM = geometric mean. a. Analyses of the same birth cohort recruited in the Sirajdikhan and Pabna Sadar upazilas of Bangladesh.

Table 2.3: Summary of epidemiological studies of folate and arsenic methylation.

Study	**Location**	**Design**	**N**	**Measure of folate**	**Results related to folate[a]**	**Results related to other OCM micronutrients[a]**
Basu et al. (2011); Mitra et al. (2004)	West Bengal, India	CC	405	24-h dietary recall; serum folate	Quintile of folate intake not associated with skin lesions. Inverse trend in OR for skin lesions with quintile of folate intake. Lowest tercile of serum folate associated with lower urine %InAs and higher %MMAs compared to the highest.	Lowest quintile of animal protein intake associated with skin lesions. Inverse trend in OR for skin lesions with quintile of animal protein intake. Lowest tercile of urine creatinine associated with higher urine %InAs, and lower %MMAs and %DMAs compared to highest. Lowest tercile of riboflavin intake associated with lower urine %InAs and higher %DMAs compared to highest. Lowest tercile of plasma Hcy associated with lower urine %MMAs compared to highest. Lowest tercile of animal protein intake associated with lower urine %MMAs compared to highest.
Steinmaus et al. (2014)	Western US	CS	87	Dietary questionnaire	Folate intake not associated with urine As metabolite proportions	Lowest quartile of protein intake associated with higher urine %MMAs and lower %DMAs compared to highest.

Study	Location	Design	N	Measure of folate	Results related to folate[a]	Results related to other OCM micronutrients[a]
Gamble et al. (2005b)	Bangladesh	CS	300	Plasma folate	Plasma folate associated with lower urine %InAs and %MMAs, and higher %DMAs	Total Hcy associated with higher urine %MMAs and lower %DMAs Cysteine associated with lower urine %InAs and higher %MMAs Urinary creatinine associated with lower urine %InAs and higher %DMAs
Gamble et al. (2006, 2007)	Bangladesh	RCT	300; folate deficient	400 μg FA per day for 12 weeks	FA treatment associated with greater decrease in urine %InAs and %MMAs, and increase in %DMAs compared to placebo FA treatment associated with greater decrease in total bAs and blood MMAs concentrations compared to placebo	Urinary creatinine associated with lower urine %InAs and higher %DMAs
Li et al. (2007)	Bangladesh	CS	864 pregnant women	Plasma folate at gestational week 14	Among women with urine As >209 μg/L, lowest tercile of plasma folate associated with higher urine %InAs compared to highest tercile Lowest tercile of plasma folate, B_{12}, and zinc associated with higher urine %InAs and primary methylation index compared to highest tercile of plasma folate, B_{12}, and zinc	Urinary creatinine associated with lower %InAs, and higher %DMAs, primary methylation index, and SMI

Study	Location	Design	N	Measure of folate	Results related to folate[a]	Results related to other OCM micronutrients[a]
Pilsner et al. (2009)	Bangladesh	Nested CC	548	Plasma folate	Low folate (<9 nmol/L) associated with increased risk of skin lesions compared to high folate Compared to the referent group (low Hcy + high folate), groups with low Hcy + low folate, hyperhomocysteinemia + high folate, and hyperhomocysteinemia + low folate had increased risk of skin lesions	Hyperhomocysteinemia associated with increased risk of skin lesions compared to low homocysteine Urinary creatinine associated with decreased risk of skin lesions Highest quartiles of urinary creatinine associated with decreased risk of skin lesions compared to lowest quartile
Gardner et al. (2011)	Bangladesh	CS	324 pregnant women	Plasma folate	Plasma folate not associated with change in urine As proportions	Plasma B_{12} not associated with change in urine As proportions
Bozack et al. (2018b); Peters et al. (2015a)	Bangladesh	RCT	610	400 or 800 μg FA, 3 g creatine, or 400 μg FA + 3 g creatine/ day for 12 weeks	800 μg FA associated with greater decrease in total bAs compared to placebo FA associated with greater decrease in urine %InAs and %MMAs, and increase in %DMAs compared to placebo	Creatine associated with greater decrease in urine %MMAs compared to placebo

Study	Location	Design	N	Measure of folate	Results related to folate[a]	Results related to other OCM micronutrients[a]
López-Carrillo et al. (2016)	Northern Mexico	CS	1,027 women	Dietary questionnaire	Folate $\geq$ 400 μg/day associated with lower urine %InAs and higher DMAs/InAs	B_{12} intake associated with lower urine %InAs and higher %DMAs, SMI, and DMAs/InAs Choline and methionine intake associated with lower urine %InAs and higher %DMAs, DMAs/MMAs, and DMAs/InAs
Kurzius-Spencer et al. (2017)	US	CS	2,420 adults & children	RBC and serum folate; 24-h dietary recall	Folate intake associated with lower urine %InAs and higher %DMAs in adults	B_6 intake associated with lower urine %InAs in adults
					RBC and serum folate not associated with urine As metabolite proportions	Plasma Hcy associated with higher urine %MMAs and lower SMI in children Urinary creatinine associated with lower urine %InAs and %MMAs, and higher %DMAs and SMI
					RBC folate associated with lower urine %InAs	
					Folate intake and serum folate not associated with urine As metabolite proportions or SMI	B_{12} intake associated with higher urine %InAs and lower %DMAs B_6 intake associated with lower urine %InAs Plasma Hcy associated with higher urine %MMAs, and lower %DMAs and SMI

Study	Location	Design	N	Measure of folate	Results related to folate[a]	Results related to other OCM micronutrients[a]
Spratlen et al. (2017)	US	CS	405 American Indians	Dietary questionnaire	Folate intake not associated with urine As metabolite proportions in fully adjusted models	Highest tercile of B2 intake associated with lower urine %MMAs and higher %DMAs compared to lowest tercile
					High folate intake and high B_6 intake associated with lower urine %InAs and higher %DMAs compared to low folate intake and low B6 intake	Highest tercile of B_6 intake associated with lower urine %InAs and %MMAs, and higher %DMAs compared to lowest tercile
						First principal component of OCM nutrients (representing intake of all OCM nutrients) associated with lower urine %InAs and %MMAs, and higher %DMAs

a. All results were significant at $P < 0.05$. Abbreviations: CC, Case-control; CS, Cross-sectional; RCT, Randomized controlled trial; RBC, red blood cell; %DMAs, percent dimethyl-arsenical species; FA, folic acid; Hcy, homocysteine; %InAs, percent inorganic arsenic; %MMAs, percent monomethyl-arsenical species; OCM, one-carbon metabolism; OR, odds ratio; SMI, secondary methylation index.

Table 2.4: Studies identified reporting the association between As exposure and DNA methylation measured using the 450K or 850k microarray.

Study	**Location**	**N**	**Exposure assessment**	**Mean exposure**	**DNAm assessment**	**Platform**	**Number of significant CpGs**
Koestler et al. (2013)	US	134; birth cohort	Maternal urinary As	Mean = 4.1 µg/L	Cord blood	450K	0 after correction for multiple testing
Broberg et al. (2014)	Bangladesh	127; birth cohort	Maternal urinary As	Mean = 66 µg/L	Cord blood	450K	3 in boys ($FDR < 0.05$)
Kile et al. (2014)[a]	Bangladesh	44; birth cohort	Maternal drinking water As	Median = 12 µg/L	Cord blood	450K	1 ($FDR < 0.05$)
Liu et al. (2014)	US	46; adults	Toenail As	Low exposure: <0.0649 µg/kg; high exposure: ≥ 0.1442 µg/kg	Whole blood cells	450K	0 after correction for multiple testing
Argos et al. (2015)	Bangladesh	400; adults	Blood and urinary As	bAs, mean = 9.3 µg/L; urinary As, mean = 302 µg/g creatinine	Whole blood	450K	4 ($P_{Bonferroni} < 0.05$)
Rojas et al. (2015)	Mexico	38; birth cohort	Maternal urinary As	Mean = 73.9 µg/L	Cord blood	450K	4,771 ($FDR < 0.05$)

Study	Location	N	Exposure assessment	Mean exposure	DNA methylation assessment	Platform	Number of significant CpGs
Cardenas et al. (2015)	Bangladesh	Placenta: N = 37; cord blood: N = 45; human umbilical vein endothelial cells (HUVEC): N = 52; birth cohort	Maternal drinking water As	Mean = 63.7 µg/L	Placenta, cord blood, and HUVEC	450K	5 ($P_{Bonferroni} < 0.05$)
Green et al. (2016)	US	343; birth cohort	Maternal toenail As and placental As	Maternal toenail As, mean = 0.05 µg/kg; placental As, mean = 0.82 µg/kg	Placenta	450K	Maternal toenail arsenic: 1; placental arsenic: 163 ($FDR < 0.05$)
Kaushal et al. (2017)	Taiwan	64; birth cohort	Maternal urinary As	Mean = 23.2 µg/g creatinine	Cord blood	450K	579 ($FDR < 0.05$)
Ameer et al. (2017)	Argentina	96; women	Urinary As	Median = 185 µg/L	Whole blood	450K	6 ($FDR < 0.05$)
Bozack et al. (2018a)[a]	Bangladesh	44; birth cohort	Maternal drinking water As	Mean = 63.7 µg/L	Cord blood	450K	380 ($P < 1.10\times10^{-6}$; $\lvert\beta$ regression$\rvert > 0.10$); top 10 reported

Study	**Location**	**N**	**Exposure assessment**	**Mean exposure**	**DNA methylation assessment**	**Platform**	**Number of significant CpGs**
Guo et al. (2018)	China	102; families (adults and children)	Village drinking water As > or < 10 μg/L	Among exposed group, mean = 266.4 ppb	Buffy coat	450K	85 ($FDR < 0.05$)
Gliga et al. (2018)	Bangladesh	113; birth cohort	Maternal urinary As	Mean = 68 μg/L	Blood mononuclear cells collected at 9 years	450K	9 in all children, 57 in boys, 15 in girls ($FDR < 0.05$)
Demanelis et al. (2019)	Bangladesh	396; adults	Urinary As; drinking water As	Urinary As, median = 201.5 μg/g creatinine; water As median: 50.5 μg/L	Whole blood	850K	Urinary As: 34 ($FDR < 0.05$); water As: 24 ($FDR < 0.05$)

DMAm, DNA methylation; a. Analyses of the same birth cohort recruited in the Sirajdikhan and Pabna Sadar upazilas of Bangladesh.

Chapter 3

Folic acid supplementation enhances arsenic methylation: results from a folic acid and creatine supplementation randomized controlled trial in Bangladesh

Anne K. Bozack,[1] Megan N. Hall,[2] Xinhua Liu,[3] Vesna Ilievski,[1] Angela M. Lomax-Luu,[1] Faruque Parvez,[1] Abu B. Siddique,[4] Hasan Shahriar,[4] Mohammad N. Uddin,[4] Tariqul Islam,[4] Joseph H. Graziano,[1] Mary V. Gamble[1]

Affiliations: 1. Dept. of Environmental Health Sciences, Columbia University; 2. Dept. of Epidemiology, Columbia University; 3. Dept. of Biostatistics, Columbia University; 4. Columbia University Arsenic Project in Bangladesh, Dhaka, Bangladesh

Published: https://www.ncbi.nlm.nih.gov/pubmed/30590411

Bozack et al. Folic acid supplementation enhances arsenic methylation: results from a folic acid and creatine supplementation randomized controlled trial in Bangladesh. *Am J Clin Nutr.* (2019), doi: 10.1093/ajcn/nqy148.

3.1 Abstract

Background: Arsenic (As) exposure through drinking water persists in many regions. Inorganic As (InAs) is methylated to monomethyl-arsenical species (MMAs) and dimethyl-arsenical species (DMAs), facilitating urinary excretion. Arsenic methylation is dependent on one-carbon metabolism, which is influenced by nutritional factors such as folate and creatine.

Objective: This study investigated the effects of folic acid (FA) and/or creatine supplementation on the proportion of As metabolites in urine.

Design: In a 24-week randomized, double-blinded, placebo-controlled trial, 622 participants were assigned to receive FA (400 or 800 μg/day), 3 g creatine/day, 400 μg FA + 3 g creatine/day, or placebo. The majority of participants were folate sufficient; all received As-removal water filters. From weeks 12-24, half of the participants receiving FA received placebo.

Results: Among groups receiving FA, the mean decrease in ln(%InAs) and %MMAs and increase in %DMAs exceeded that of the placebo group at weeks 6 and 12 ($P < 0.05$). In the creatine group, the mean decrease in %MMAs exceeded that of the placebo group at weeks 6 and 12 ($P < 0.05$); creatine supplementation did not affect change in %InAs or %DMAs. The decrease in %MMAs at weeks 6 and 12 was larger in the 800 μg FA than the 400 μg FA group ($P = 0.034$). There were no differences in treatment effects between the 400 μg FA and creatine+FA groups. Data suggest a rebound in As metabolite proportions after FA cessation; at week 24, ln(%InAs) and %DMAs were not significantly different than baseline levels among participants who discontinued FA supplementation.

Conclusions: The results of this study confirm that FA supplementation rapidly and significantly increases methylation of InAs to DMAs. Further research is needed to understand the strong cross-sectional associations between urinary creatinine

and As methylation in previous studies.

Acknowledgments

We thank and our staff, the fieldworkers, and the study participants in Bangladesh, without whom this work would not have been possible.

3.2 Introduction

Chronic arsenic (As) exposure is a major global public health issue (Naujokas et al., 2013). In Bangladesh, roughly 57 million people are exposed to concentrations in drinking water exceeding the World Health Organizations guideline of 10 μg/L (Government of the People's Republic of Bangladesh et al., 2001; World Health Organization, 2012). Chronic As exposure is associated with a range of health outcomes including skin lesions, skin and internal cancers, neurological impairment, cardiovascular and pulmonary diseases, and endocrine disruption (Naujokas et al., 2013).

Ingested inorganic As (InAs) undergoes a series of stepwise biotransformation reactions in which it is methylated by arsenic methyltransferase (AS3MT) (Thomas et al., 2007) using the methyl donor *S*-adenosylmethionine (SAM) to form monomethylarsonic acid ($MMAs^V$) and dimethylarsinic acid ($DMAs^V$) (Figure 3.1) (Challenger, 1945). Variation in As methylation capacity between populations and individuals may modify the association between As exposure and health outcomes (Loffredo et al., 2003). Complete methylation of As to $DMAs^V$ is important as toxicological studies have demonstrated that $MMAs^{III}$ is the most toxic As species (Petrick et al., 2000; Moe et al., 2016). Although the relatively toxicity of $MMAs^{III}$ and $MMAs^V$ is difficult to assess in human studies because $MMAs^{III}$ is rapidly oxidized to $MMAs^V$, a higher proportion of urinary $MMAs^{(III+V)}$ has been associated with higher risk of As-induced skin lesions (Zhang et al., 2014; Yu et al., 2000; Pierce et al., 2013), peripheral vascular disease, skin and bladder cancer, and atherosclerosis, as reviewed by Steinmaus et al. 2010.

Arsenic methylation is dependent on one-carbon metabolism for SAM synthesis (Hall and Gamble, 2012). In this pathway, the one-carbon unit carried by 5-methyl-tetrahydrofolate (5-MTHF) is transferred to homocysteine to form methionine, which is activated to form SAM. Dietary creatine may also impact As metabolism. En-

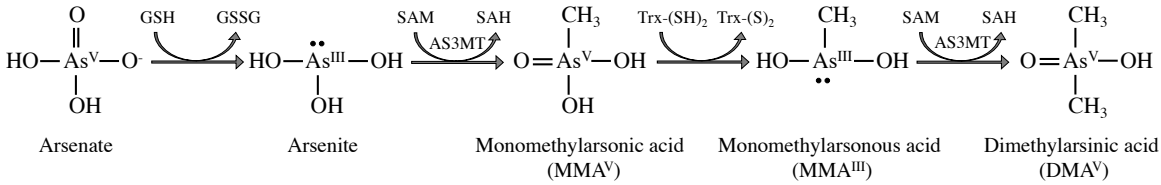


Figure 3.1: Arsenic methylation. Ingested InAs undergoes a series of stepwise biotransformation reactions. As^{III} is actively methylated by AS3MT using the methyl donor SAM to form $MMAs^{V}$. $MMAs^{V}$ is reduced to $MMAs^{III}$, and $MMAs^{III}$ is methylated to form $DMAs^{V}$. Abbreviations: As, arsenic; As^{III}, arsenite; AS3MT, arsenic arsenic (+3 oxidation state) methyltransferase; $DMAs^{V}$, dimethylarsinic acid; InAs, inorganic As; $MMAs^{V}$, monomethylarsonic acid; $MMAs^{III}$, monomethylarsonous acid; SAM, *S*-adenosylmethionine.

dogenous biosynthesis of creatine consumes an estimated 40% of SAM (Brosnan et al., 2011). Dietary creatine consumption from animal sources downregulates endogenous creatine biosynthesis, reducing the loss of methyl groups incurred by creatine biosynthesis. Among omnivores, approximately half of the bodys supply of creatine is provided by food (Brosnan et al., 2011).

We have examined the role of folate nutritional status on As methylation in several prior studies (Gamble et al., 2005, 2006, 2007). The effects of supplementation with 400 μg folic acid (FA)/day, the U.S. recommended daily allowance (RDA), among Bangladeshi adults with low plasma folate (< 9 nmol/L) were tested in a randomized controlled trial (RCT). After 12 weeks, the treatment group had a significantly larger increase in the percentage of urinary dimethyl-arsenical species (%DMAs) and decreases in urinary log(%InAs), the percentage of urinary monomethyl-arsenical species (%MMAs), (Gamble et al., 2007) total blood As (bAs), and blood MMAs (Gamble et al., 2007) as compared to the placebo group. Changes in total bAs in the RCT we present here have been previously published (Peters et al., 2015a). Among As-exposed Bangladeshi adults selected independent of folate status, a larger decrease in ln(bAs) was observed in the treatment group receiving 800 μg FA/day for 12 weeks relative to the placebo group. Urinary creatinine, a product of creatine, has also been negatively

associated with urinary %InAs and positively associated with urinary %DMAs (Gamble et al., 2006, 2005; Ahsan et al., 2007; Hall et al., 2009; Basu et al., 2011; Peters et al., 2014).

It is not known if FA supplementation increases As methylation in a population of mixed folate-deficient and -sufficient individuals, which may have important policy implications. The current RCT investigated the effects of 400 μg/day and 800 μg/day FA supplementation on urinary As metabolites among the general population of As-exposed adults in Araihazar, Bangladesh. We also tested the effect of reducing methylation demand through creatine supplementation on As methylation.

3.3 Subjects

The Folic Acid and Creatine Trial (FACT) is an RCT conducted in 2010-2012 that has been previously described (Peters et al., 2015a). Briefly, FACT participants were recruited from the Health Effects of Arsenic Longitudinal Study (HEALS), a prospective cohort study of > 30,000 adults living in Araihazar, Bangladesh that was initiated in 2000 (Ahsan et al., 2006). HEALS participants were married, residing in the study area for at least five years, and drinking from their household well for at least three years. FACT participants were randomly selected from the HEALS cohort and were eligible if they were drinking for at least one year from a household well that had an As concentration $\geq$ 50 μg/L. Exclusion criteria were pregnancy, taking nutritional supplements, having proteinuria, renal disease, diabetes, gastrointestinal problems, or other health problems.

3.4 Methods

Study Design

The study design has been described in detail by Peters et al. (2015a). A total of 622 participants were recruited (Appendix Figure 3.5). The sample size was selected to achieve 80% power at alpha = 0.05 to detect a moderate effect size in the mean differences in change in total bAs concentrations or percent urinary As metabolites between two treatment groups (i.e., 0.45 SD) and assuming some participants might not complete the study. Participants were randomly assigned to one of five treatment groups: placebo (N = 104), 400 μg FA/day (referred to as 400FA; N = 156), 800 μg FA/day (800FA; N = 154), 3 g creatine/day (creatine; N = 104), and 3 g creatine and 400 μg FA/day (creatine+400FA; N = 104). To ensure an equal number of men and women in each treatment group and balance between treatment groups, men and women were randomized separately and in blocks, and assigned to treatment groups in the ratio 1 (placebo):1.5 (400FA):1.5 (800FA):1 (creatine):1 (creatine+400FA). The order of treatment assignments within the blocks was determined by a random permutation generated by a statistician at Columbia University. One data management specialist in the United States and one in Bangladesh assigned letters to treatment groups (i.e., A, B, C, D, E). The creatine dose of 3 g/d was selected to over-compensate for daily creatine loss, approximately 1.9 g for a 20- to 39-year-old 70-kg male, and thus to be sufficient to downregulate endogenous creatine synthesis (Brosnan et al., 2011). All supplements were obtained from Atrium Innovations, Inc. (Westmount, Quebec). Field staff received barcode-labeled pill bottles from a pharmacist and distributed bottles in order of the random treatment assignment. With the exception of the data management specialists who assigned treatment groups, all participants, field staff, village health workers, laboratory technicians, and investigators were blinded to the treatment for

the duration of the study.

The study consisted of two 12-week phases (Figure 3.2). During the first phase, participants received daily supplements or placebo according to their treatment group assignment. The primary outcome of interest (previously published) was total bAs (Peters et al., 2015a). For the current analyses, data from the first phase of the study was employed to evaluate our a priori secondary outcomes, i.e., the effects of FA and/or creatine supplementation on As methylation capacity. During the second phase of the study, participants in the 400FA and 800FA groups were randomly assigned to continue FA supplementation (400 μg FA/day, referred to as 400FA/FA: N = 77; 800 μg FA/day, 800FA/FA: N = 77) or receive placebo (400FA/placebo: N = 76; 800FA/placebo: N = 74) for the duration of the trial (weeks 12 to 24). Participants in the creatine and creatine+400FA groups were also given placebo during the second phase in order to maintain the study blind. Data from the second phase of the study was used to determine if a rebound in urinary As metabolites occurred after FA supplementation was discontinued.

Ethics

The study protocol was approved by the Columbia University Medical Center Institutional Review Board and the Bangladesh Medical Research Council. Staff physicians in Bangladesh obtained informed consent from participants.

Field work

Five trained teams, each consisting of one interviewer and one physician, recruited participants and performed all follow-up home visits during which biological samples were collected. Venous blood samples were collected at baseline, week 12, and week 24. Urine samples were collected at baseline and weeks 1, 6, 12, 13, 18, and

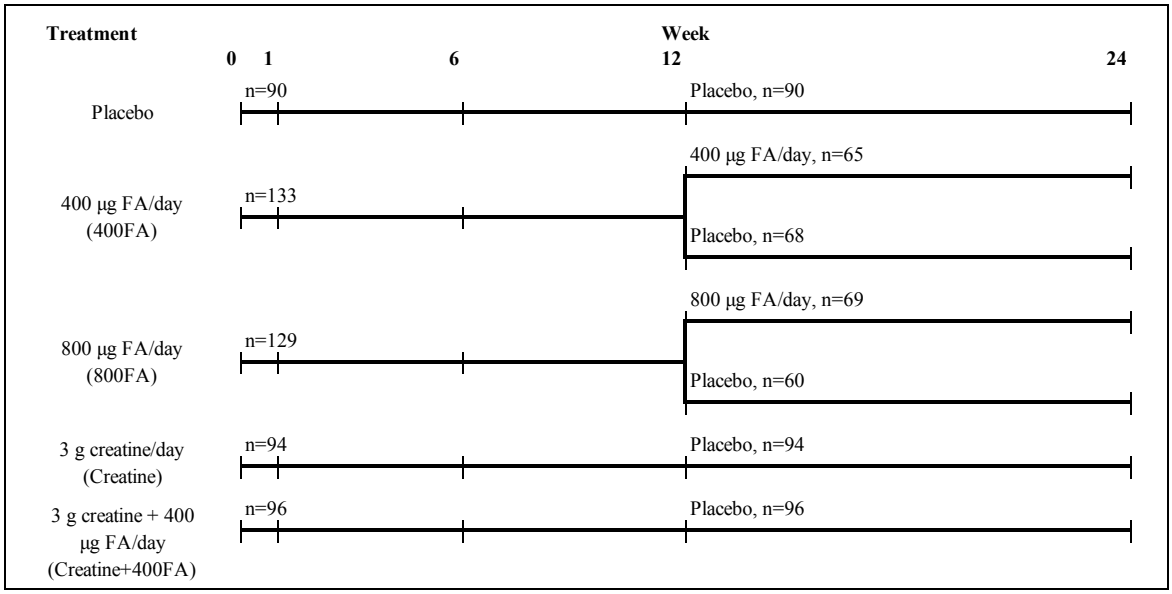


Figure 3.2: Study design. Abbreviations: 400FA, 400 µg FA/day treatment group; 800FA, 800 µg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 µg FA/day treatment group.

24. Study recruitment and follow-up was conducted from December 2009 through May 2011.

All participants received READ-F As removal filters (READ-F filter; Brota Services International, Bangladesh) and were repeatedly encouraged to use them for all drinking and cooking water throughout the study (Sanchez et al., 2016). During the first phase, all participants received two pill bottles: one containing FA pills or matched placebo pills, and one containing creatine pills or matched placebo pills. During the second phase, all participants received one pill bottle containing FA pills or matched placebo pills. Village health workers performed daily home visits, during which they observed participants take the pills or inquired about compliance. At the end of each study phase, pill bottles were collected and pill counts were performed. Compliance ranged from 79.1% to 100% (median: 99.5%; interquartile range: 98.3%, 100.0%), and did not differ substantially between treatment groups or study phases (Peters et al., 2015a).

Twelve participants were dropped from the study due to adverse events (N = 6; placebo group: abdominal cramps; 400FA group: hypertension; 800 FA group: abdominal cramps, severe vertigo, bilateral hydronephrosis; creatine group: severe vertigo), pregnancy (N = 3; 400FA, creatine, and creatine+400FA groups), missing baseline blood sample (N = 1; creatine group), and dropout (N = 2; placebo and 400FA groups). The remaining sample consisted of 610 participants (placebo: N = 102; 400FA: N = 153; 800FA: N = 151; creatine: N = 101; creatine+400FA: N = 103) (Appendix Figure 3.5).

Sample handling

Venous blood was collected in EDTA vacutainer tubes. Blood samples were stored in IsoRack cool packs (Brinkmann Instruments, Riverview, FL) at 0°C. Within four hours, samples were transported to the field clinic in Araihazar. Plasma was separated from venous blood by centrifugation at 3,000×g for 10 minutes at 4°C. Whole blood and plasma samples were stored at -80°C and shipped on dry ice to Columbia University for analysis. Urine was collected in 50-mL acid-washed polypropylene tubes, stored in portable coolers for transport to the field clinic within four hours, frozen at -20°C, and shipped to Columbia University on dry ice.

Water arsenic

To assess baseline As exposure, unfiltered water samples from the tube wells used by study participants were collected. Samples were collected in 20-mL polyethylene scintillation vials, acidified to 1% using high-purity Optima HCl (Fisher Scientific) for at least 48 hours, diluted 1:10, and analyzed using high-resolution inductively coupled plasma mass spectrometry with germanium spike to correct for fluctuations in sensitivity. To monitor As removal by the filters, filtered water was tested in the field

with Hach EZ Arsenic test kits; approximately 50 filters were replaced during the study (Sanchez et al., 2016).

Laboratory measures

Laboratory methods have been described in detail by Peters et al. (Peters et al., 2015a). Briefly, total bAs in whole blood was assessed using a PerkinElmer Elan DRC II ICP-MS (Waltham, MA) with an AS 93+ autosampler (intra- and inter-assay CVs: 2.7% and 5.7%, respectively). Plasma folate and B_{12} were assessed using radioimmunoassay (SimulTRAC-SNB; MP Biomedicals, Santa Ana, CA) (plasma folate intra- and inter-assay CVs: 5% and 13%; B12 intra- and inter-assay CVs: 6% and 17%). Plasma total homocysteine was measured with high performance liquid chromatography (HPLC) with fluorescence detection (homocysteine intra- and inter-assay CVs: 5% and 7%).

Urinary As metabolites were measured using HPLC separation of arsenobetaine, arsenocholine, As^{III}, As^{V}, MMAs ($MMAs^{III}$ + $MMAs^{V}$), and DMAs, and detection by ICP-MS with dynamic reaction cell as described by Reuter et al. 2003 (Reuter et al., 2003) (intra- and inter-assay CVs for arsenobetaine + arsenocholine; 10.1%, 12.2%; As^{III} + As^{V}: 2.7%, 4.7%; MMAs: 2.8%, 3.9%; DMAs: 0.6%, 1.3%). As^{III} can oxidize to As^{V} during sample storage and analysis, and therefore total InAs (As^{III} + As^{V}) is reported. Specific gravity (SG) was measured by refractometer.

Study sample

Six follow-up urine samples were missing, and 18 urine samples with missing SG were removed from the dataset. We suspect that some participants may have added water to urine collection cups in an effort to increase fluid volume. Therefore, urine samples with SG ≤ 1.001, outside the normal range (28), were removed from the dataset

(N = 620 out of 4239 samples). Samples with %InAs $\geq$ 55 were also considered to be outside of normal values (Vahter, 1999) and removed (N = 22 out of 3619 samples); participants without a baseline urine sample (N = 50) or without the majority of urine samples (i.e., four or more time points) (N = 17) were also removed. One participant with baseline SG of 1.002, but biologically implausibly high total urinary As, was also removed. A total of 542 participants were included in analyses (placebo: N = 90; 400FA: N = 133; 800FA: N = 129; creatine: N = 94; creatine+400FA: N = 96) (Appendix Figure 3.5). With the exception of urine As metabolites, baseline characteristics did not differ significantly between participants included and excluded from analyses.

The concentrations of urinary As metabolites were adjusted for between-individual differences in urine dilution at each time point by multiplying the sample concentration by the ratio (mean SG - 1)/(subjects SG - 1) (Miller et al., 2004). Arsenobetaine and arsenocholine are non-toxic dietary sources of As, (Syracuse Research Corporation, 2007) and were therefore excluded from analyses. Seven samples had urinary As^{III} concentrations below the limit of detection (LOD). These samples were replaced using values of LOD/2. The proportions of each As metabolite in urine were calculated by dividing the concentration of each species by the concentration of total urinary As (As^{III} + As^{V} + MMAs + DMAs).

Statistical analysis

An intent-to-treat approach was used in all analyses. Summary statistics for baseline characteristics by treatment group were calculated. Overall differences between treatment groups at baseline were detected using the Kruskall-Wallis test for continuous variables and the Chi-square test for categorical variables. One-way ANOVA was used to detect treatment group differences in urinary As metabolites and nutritional factors at baseline and follow-up, where %InAs, red blood cell (RBC) folate, plasma folate,

and homocysteine were log transformed to meet model assumptions. To compare each of the four treatment groups vs. the placebo group, Dunnett's t-test with multiple comparisons was used to control the Type I experiment-wise error rate for each As metabolite variable.

Linear models with repeated measures were applied to the outcomes of percent urinary As metabolites measured over time to examine treatment effects during the first phase. Percent InAs was log transformed to improve model fitting and reduce the impact of extreme values. Models included predictors of time, treatment group, and group by time interactions. The interaction parameters indicate treatment group differences in mean within-person change since baseline. A generalized estimating equation approach was used to estimate model parameters and test hypotheses using all available data and accounting for within-subject correlations in the repeated measures. Wald test was used to detect differences in the within-person change in As metabolite proportions between each treatment group and the placebo group. Wald test was also used to examine if within-person change in the proportion of urinary As metabolites differed between the 400FA, 800FA, and creatine+400FA treatment groups. To aid interpretation, group differences in the mean within-person changes were derived from relevant model parameters.

The rebound in urinary As metabolites after cessation of FA supplementation during the second phase (weeks 12 through 24) was similarly assessed using linear models with repeated measures applied to the 400FA and 800FA groups. The models that described the data well included parameters for treatment groups, time (weeks 0, 12, and 24), continuation status of FA supplementation, and two-way interactions of group by FA continuation, group by time, and time by FA continuation. Time by FA continuation interactions indicate differences in mean within-subject change in percent As metabolites between groups that continued and discontinued FA supplementation.

From the model parameters, we derived the differences between FA groups in the mean within-person changes over periods of interest to describe the rebound effect.

Analyses were performed using SAS 9.4 (Cary, NC) or R 3.2.2 (Vienna, Austria) (R Core Team, 2015).

3.5 Results

Baseline characteristics by treatment group are presented in Table 3.1. Participants ranged in age from 24 to 55, and approximately half (51.5%) were male. There were no statistically significant differences between treatment groups in age; sex; BMI; history of betel nut use or smoking; land ownership; water, blood, and urinary As concentrations; urinary As species concentrations; or nutritional status, including red blood cell and plasma folate, folate status, plasma B_{12}, B_{12} status, plasma homocysteine, and hyperhomocysteinemia. Urinary creatinine was significantly correlated with %InAs at baseline (Spearman correlation coefficient r_s = -0.28, $P < 0.001$) and %DMAs (r_s = 0.13, $P = 0.002$). There were no significant treatment group differences in baseline percent urinary As metabolites (Kruskall-Wallis test %InAs $P = 0.136$; %MMAs $P = 0.578$; %DMAs $P = 0.235$).

After treatment began, there was an initial decrease in total urinary As among all treatment groups as a result of using As removal filters. There was also a notable initial decrease in SG and urinary creatinine concentrations, suggesting that participants were drinking more water after receiving the filters. We consider this to be a new toy effect, however, as total urinary As concentration, SG, and urinary creatinine concentration gradually increased after the first week, indicating that participants water intake returned to normal and that they used the water filters less frequently as the study progressed. After adjusting for SG, the trend in total urinary As remained (Figure 3.3). Use of the water filters and total urinary As concentrations during this

Table 3.1: Baseline characteristics

Characteristic	**Placebo (N=90)** Mean ± SD	**400FA (N=133)** Mean ± SD	**800FA (N=129)** Mean ± SD	**Creatine (N=94)** Mean ± SD	**Creatine+400FA (N=96)** Mean ± SD	***P*[1]**
Age (years)	37.5 ± 7.2	39.3 ± 8.2	38.4 ± 8.2	38.5 ± 8.3	38.0 ± 7.8	0.532
Male (%)	48.9	53.4	51.9	51.1	51.0	0.977
Smoking ever (%)[2]	25.6	23.7	31.8	27.7	30.2	0.62
Betel nut use ever (%)[3]	25.6	25.2	25.6	24.5	20.8	0.929
Owns land (%)	46.7	51.1	45	47.9	41.1	0.651
BMI [4]	20.5 ± 3.1	19.5 ± 2.3	19.8 ± 2.7	20.0 ± 2.9	19.5 ± 2.5	0.172
Water As (μg/L)	151.4 ± 120.6	146.8 ± 118.7	156.9 ± 128	153 ± 118.4	170.2 ± 150.5	0.92
Urinary As (μg/L)	171.1 ± 115.4	190.2 ± 184	177.1 ± 113.9	188.2 ± 150.1	187.8 ± 107.2	0.383
Blood As (μg/L)	9.6 ± 5.3	10.2 ± 7.7	9.9 ± 5.4	10.6 ± 8.2	10.6 ± 5.3	0.32
Urinary InAs (μg/L)	23.2 ± 18.4	24.6 ± 28.7	22.5 ± 18.2	26.9 ± 28.1	22.1 ± 12.2	0.532
Urinary MMAs (μg/L)	20.7 ± 19.1	23.1 ± 29.2	22.3 ± 21.9	23.7 ± 21.6	21.7 ± 14.2	0.611
Urinary DMAs (μg/L)	109.9 ± 76.3	118.2 ± 99.2	116.6 ± 73.2	117.8 ± 80.8	124.3 ± 66.1	0.156
Urinary %InAs	14.7 ± 4.2	13.9 ± 4.3	13.8 ± 4.5	14.6 ± 5.3	13.2 ± 3.7	0.136
Urinary %MMAs	13 ± 4.2	13.1 ± 4.5	13.1 ± 5	13.4 ± 4.3	12.5 ± 4.3	0.578
Urinary %DMAs	72.4 ± 7	73 ± 6.4	73.1 ± 7.6	72 ± 7.5	74.4 ± 6.1	0.235
Urinary creatinine (mg/dL)	52.4 ± 35.3	61.2 ± 44.8	57.3 ± 41.1	59.2 ± 36.2	66.1 ± 46.8	0.346
RBC folate (nmol/L)[4]	476 ± 187.9	509.5 ± 351.8	491.6 ± 174.6	-	-	0.772
Plasma folate (nmol/L)[5]	16.8 ± 18.2	16.9 ± 14.8	18.1 ± 16.8	16.0 ± 7.6	15.3 ± 8.8	0.625
Folate deficient (%)[5,6]	23.3	23.3	17.8	12.9	20.8	0.298
Plasma B12 (pmol/L)	221 ± 99.4	247.5 ± 135	250.9 ± 146.9	256.7 ± 145.8	235.5 ± 117.9	0.739
B12 deficient (%)[7]	26.7	25.6	26.4	21.5	24.0	0.918
Plasma homocysteine (μmol/L)	13.1 ± 5.7	13.8 ± 9.2	14 ± 10.9	12.2 ± 5.6	12.7 ± 5.5	0.79
Hyperhomocysteinemia (%)[8]	43.3	36.8	41.1	37.6	38.5	0.87

1. Kruskall-Wallis test for continuous variables and chi-squared test for categorical variables. 2. FA 400: N = 131. 3. Placebo: N = 89; 400FA: N = 130; 800FA: N = 126; Creatine: N = 92; Creatine + 400FA: N = 95. 4. Placebo: N = 89, 400FA: N = 130; 800FA: N = 128; Not measured for Creatine and Creatine + 400 FA groups. 5. Creatine: N = 93. 6. $<$ 9 nmol/L in plasma. 7. $<$ 151 pmol/L. 8. $\geq$ 13 μmol/L.

Abbreviations: FA, folic acid; 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group, arsenic (As), InAs, inorganic arsenic; DMAs, dimethyl-arsenical species; MMAs, monomethyl-arsenical species; RBC, red blood cell.

trial has been detailed by Sanchez et al 2016. At week 12, the mean total urinary As concentration was 15% lower than baseline among all participants ($P < 0.0001$). At a re-visit conducted one year after the study was completed, the majority of participants reported decreasing or stopping use of the filters because the water was filtered more slowly over time.

Treatment effects during the first phase

As we have previously reported, FA supplementation increased plasma and RBC folate (Peters et al., 2015a) and decreased homocysteine (Peters et al., 2015b) at week 12 (Table 3.2). Table 3.2 and Figure 3.4 present the mean percent urinary As metabolites at baseline and weeks 1, 6, and 12. Overall treatment group differences in log(%InAs) were significant at weeks 1, 6, and 12; differences in %MMAs and %DMAs were significant at weeks 6 and 12 (ANOVA $P < 0.05$). At week 1, the 800FA group had significantly lower log(%InAs) than the placebo group (Dunnett's t-test $P < 0.05$). At weeks 6 and 12, the three groups receiving FA supplements had significantly lower %InAs and %MMAs, and higher %DMAs compared to the placebo group (Dunnett's t-test $P < 0.05$). The proportions of urinary InAs, MMAs, and DMAs were not significantly different between the creatine and placebo groups at any follow-up (Dunnett's t-test $P > 0.05$).

Treatment group differences in within-person change in urinary As metabolites between baseline and weeks 1, 6, and 12 were examined using available data from the 542 subjects included 1,966 observations (week 0: N = 542; week 1: N = 456; week 6: N = 461; week 12: N = 507). The treatment effects on within-person changes in As metabolite proportions over the first 6 weeks were similar to the effects on the changes over 12 weeks, indicated by a linear model with 20 parameters where the time variable had four categories (week 0, week 1, week 6, week12). Therefore, data were described

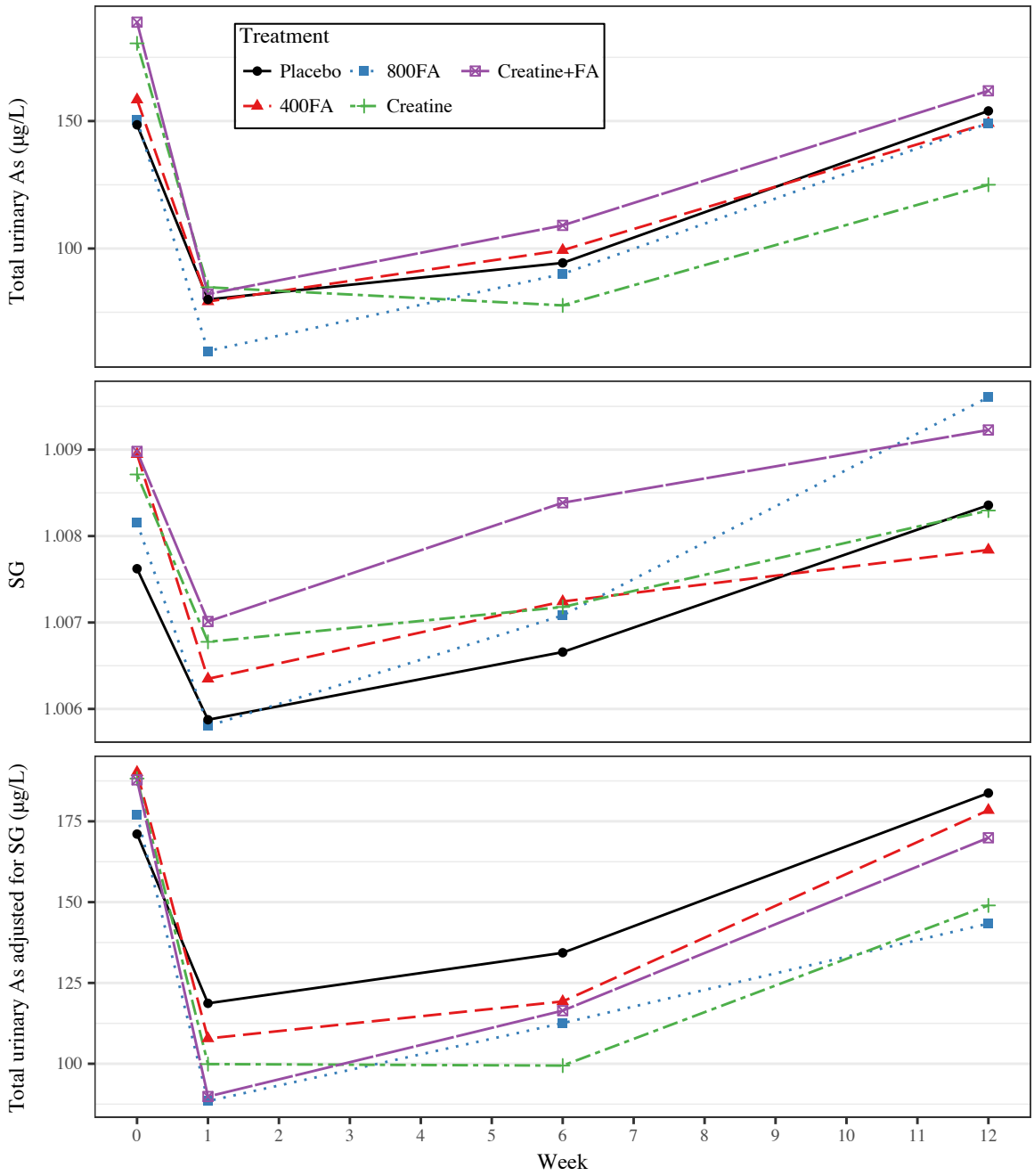


Figure 3.3: Urinary arsenic (As) concentrations over time. Total urinary As, SG, and total urinary As adjusted for specific gravity. Placebo (N = 90); 400FA (N = 133); 800FA (N = 129); Creatine (N = 94); Creatine+400FA (N = 96). Abbreviations: Abbreviations: 400FA, 400 µg FA/day treatment group; 800FA, 800 µg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 µg FA/day treatment group; As, arsenic; SG, specific gravity.

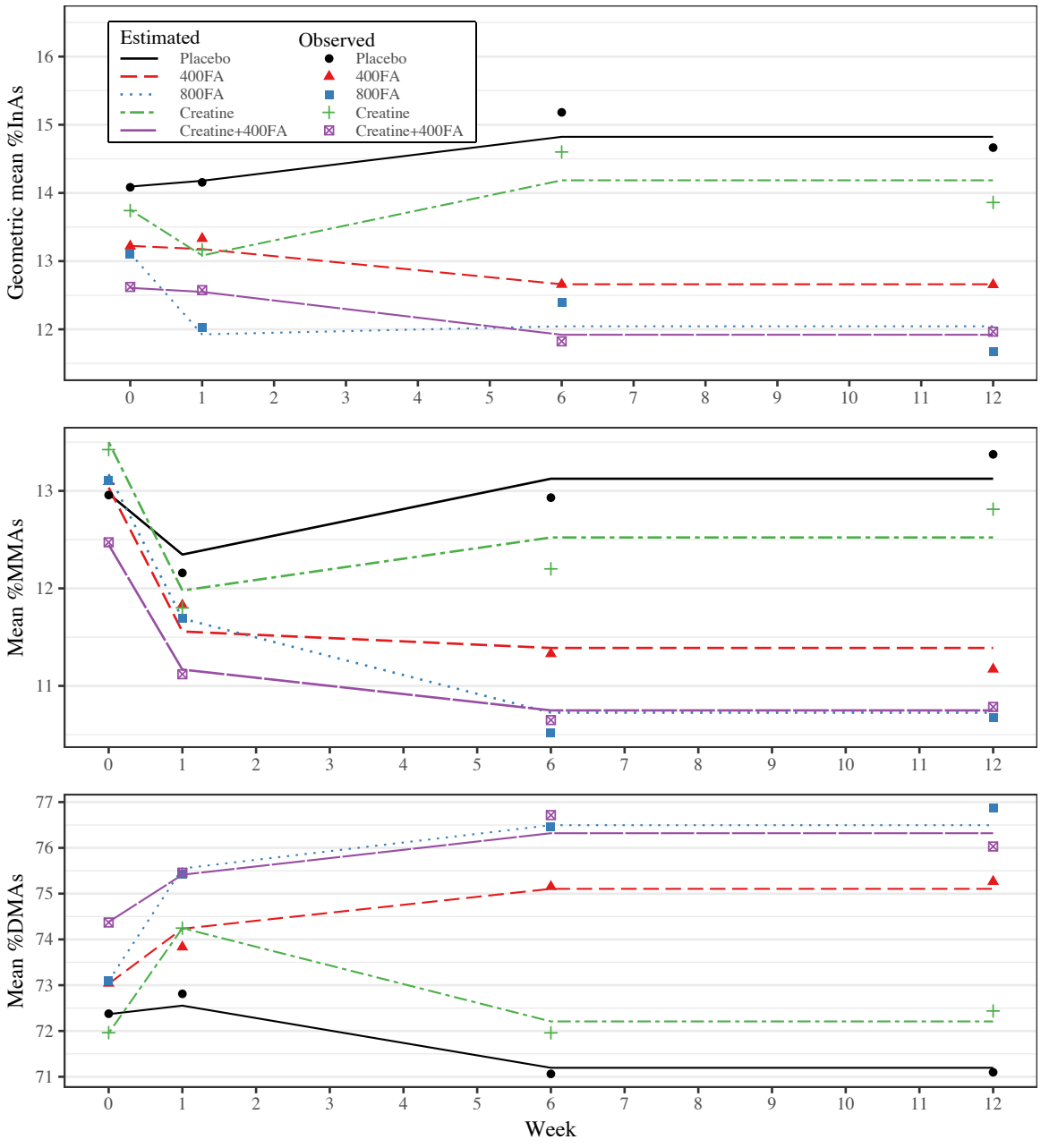


Figure 3.4: (Caption next page.)

Figure 3.4: (Previous page.) Percent urinary As metabolites at weeks 0-12. Symbols represent raw means, and lines represent the predicted means from a linear model with repeated measures of urinary As metabolites. As shown in Table 2, log(%InAs) was significantly different between treatment groups at weeks 1, 6, and 12; %MMAs and %DMAs were significantly different between treatment groups at weeks 6 and 12 (ANOVA P < 0.001). The 800FA group had significantly lower log(%InAs) at week 1 than the placebo group (Dunnett's t-test $P < 0.05$). The three groups receiving FA supplements had significantly lower %InAs and %MMAs, and higher %DMAs than the placebo group at weeks 6 and 12 (Dunnett's t-test $P < 0.05$). The proportions of urinary As metabolites were not significantly different between the creatine and placebo groups at any follow-up (Dunnett's t-test $P < 0.05$). Placebo (N = 90); 400FA (N = 133); 800FA (N = 129); Creatine (N = 94); Creatine+400FA (N = 96). Abbreviations: Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group; As, arsenic; InAs, inorganic arsenic; DMAs, dimethyl-arsenical species; MMAs, monomethyl-arsenical species.

using a parsimonious model with three categories for the time variable (week 0, week 1, weeks 6 and 12). Table 3.3 displays results from the parsimonious model. The observed geometric mean of %InAs and means of %MMAs and %DMAs were similar to the model-based estimates (Figure 3.4), indicating that the model fit the data well. There was a significantly greater mean decrease in mean ln(%InAs) from baseline to week 12 in the 400FA, 800FA, and creatine+400FA groups than the placebo group ($P < 0.05$). There were no significant group differences in the mean change in ln(%InAs) between the creatine and placebo groups. The mean decrease in %MMAs at week 1 and weeks 6 and 12 was significantly greater than the placebo group for all treatment groups ($P < 0.05$) except for the creatine+400FA group at week 1. The mean within-person increase in %DMAs was significantly larger in the 800FA group than the placebo group at week 1; and was significantly larger in the 400FA, 800FA, and creatine+400FA groups than the placebo group at weeks 6 and 12 ($P < 0.05$). The mean increase in %DMAs at weeks 6 and 12 was not significantly different between the creatine group and the placebo group.

Using linear models with repeated measures, we also examined if within-

Table 3.2: Arsenic metabolites and nutritional factors by treatment group during the first phase (baseline - week 12)

		Placebo (N=90) Mean ± SD	**400FA (N=133) Mean ± SD**	**800FA (N=129) Mean ± SD**	**Creatine (N=94) Mean ± SD**	**Creatine+400FA (N=96) Mean ± SD**	P[1]
%InAs	Baseline	14.7 ± 4.2	13.9 ± 4.3	13.8 ± 4.5	14.6 ± 5.3	13.2 ± 3.7	0.145
	Week 13[2]	15.0 ± 5.8	14.3 ± 5.9	12.9 ± 4.9*	14 ± 4.8	13.4 ± 5.0	0.038
	Week 6[3]	16.0 ± 5.6	13.5 ± 5.1*	13 ± 4.2*	15.8 ± 6.8	12.6 ± 4.8*	<0.001
	Week 12[4]	15.5 ± 5.5	13.6 ± 5.6*	12.4 ± 4.7*	14.8 ± 5.5	13.2 ± 7.0*	<0.001
%MMAs	Baseline	13.0 ± 4.2	13.1 ± 4.5	13.1 ± 5	13.4 ± 4.3	12.5 ± 4.3	0.578
	Week 1[2]	12.2 ± 3.8	11.8 ± 4	11.7 ± 4.4	11.8 ± 3.6	11.1 ± 3.8	0.522
	Week 6[3]	12.9 ± 4.3	11.3 ± 4.1*	10.5 ± 3.8*	12.2 ± 4	10.6 ± 3.7*	<0.001
	Week 12[4]	13.4 ± 3.6	11.2 ± 4.1*	10.7 4.0*	12.8 ± 4	10.8 ± 4.1*	<0.001
%DMAs	Baseline	72.4 ± 7.0	73.0 6.4	73.1 ± 7.6	72.0 ± 7.5	74.4 ± 6.1	0.159
	Week 1[2]	72.8 ± 7.3	73.8 ± 7.3	75.4 ± 7.4	74.2 ± 6.6	75.5 ± 6.9	0.073
	Week 6[3]	71.1 ± 7.7	75.2 ± 7.0*	76.5 ± 6.0*	72 ± 8.1	76.7 ± 6.5*	<0.001
	Week 12[4]	71.1 ± 7.5	75.3 ± 7.0*	76.9 ± 6.3*	72.4 ± 7.6	76.0 ± 7.8*	<0.001
RBC folate (nmol/L)[5]	Baseline[6]	476.0 ± 187.9	509.5 ± 351.8	491.6 ± 174.6	-	-	0.727
	Week 12[7]	564.9 ± 264.2	1018.8 ± 365.8*	1183.4 ± 412.3*	-	-	<0.001
Plasma folate (nmol/L)	Baseline 8[8]	16.8 ± 18.2	16.9 ± 14.8	18.1 ± 16.8	16 ± 7.6	15.3 ± 8.8	0.683
	Week 12[9]	26.8 ± 43.1	52.8 ± 45.2*	86.8 ± 156.5*	41.5 ± 101.1	81.1 ± 144.1*	<0.001
Plasma homocysteine	Baseline	13.1 ± 5.7	13.8 ± 9.2	14 ± 10.9	12.2 ± 5.6	12.7 ± 5.5	0.712
(μmol/L)	Week 12[10]	± 5.9	9.8 ± 3.3*	9.6 ± 4.1*	11.5 ± 4.2	9.7 ± 2.8*	<0.001

1. ANOVA for treatment group differences; %InAs, red blood cell folate, plasma folate, and homocysteine were log transformed to meet model assumptions. 2. Placebo: N = 80; 400FA: N = 109; 800FA: N = 96; Creatine: N = 85; Creatine+400FA: N = 86. 3. Placebo: N = 79; 400FA: N = 111; 800FA: N = 110; Creatine: N = 78; Creatine+400FA: N = 83. 4. Placebo: N = 84; 400FA: N = 125; 800FA: N = 122; Creatine: N = 88; Creatine+400FA: N = 88. 5. Not measured for the Creatine and Creatine+400FA groups. 6. Placebo: N = 89; 400FA: N = 130; 800FA: N = 128. 7. Placebo: N = 84; 400FA: n=123; 800FA: N = 122. 8. Creatine: N = 93. 9. Placebo: N = 84; 400FA:N = 125; 800FA: N = 121; Creatine: N = 88; Creatine+400FA: N = 88. 10. Placebo: N = 84; 400FA: N = 125; 800FA: N = 122; Creatine: N = 87; Creatine+400FA: N = 88.

* Dunnett's t-test with multiple comparisons for each treatment group vs. the placebo group, $P<0.05$.

Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group; As, arsenic; InAs, inorganic arsenic, DMAs, dimethyl-arsenical species; MMAs, monomethyl-arsenical species; RBC, red blood cell.

Table 3.3: Treatment group differences in mean within-person change since baseline in As metabolite proportions[1]

		Mean change Week 1 (95% CI) (week 1 - week 0)	**Mean change Weeks 6 & 12 (95% CI) (week ≥ 6 - week 0)**
ln(%InAs)	Placebo	0.01 (-0.08, 0.09)	0.05 (0.0, 0.10)
	400FA	-0.01 (-0.12, 0.10)	-0.09 (-0.17, -0.01)*
	800FA	-0.10 (-0.22, 0.02)	-0.14 (-0.21, -0.06)***
	Creatine+400FA	-0.01 (-0.12, 0.10)	-0.11 (-0.18, -0.03)**
	Creatine	-0.06 (-0.17, 0.06)	-0.02 (-0.10, 0.06)
%MMAs	Placebo	-0.62 (-1.13, -0.12)	0.15 (-0.37, 0.68)
	400FA	-0.85 (-1.57, -0.14)*	-1.80 (-2.53, -1.07)****
	800FA	-0.86 (-1.62, -0.09)*	-2.60 (-3.35, -1.85)****
	Creatine+400FA	-0.66 (-1.43, 0.11)	-1.85 (-2.61, -1.09)****
	Creatine	-0.90 (-1.74, -0.06)*	-1.13 (-2.08, -0.19)*
%DMAs	Placebo	0.19 (-1.30, 1.67)	-1.17 (-2.18, -0.17)
	400FA	1.02 (-0.85, 2.88)	3.25 (1.81, 4.68)****
	800FA	2.27 (0.26, 4.28)*	4.57 (3.20, 5.95)****
	Creatine+400FA	0.85 (-1.04, 2.74)	3.11 (1.67, 4.55)****
	2.11 (0.01, 4.21)*	1.43 (-0.21, 3.06)	

1. Treatment group differences in mean changes were derived from relevant group by time interaction parameters of the linear models with repeated measures where time was a variable with three categories (week 0, week 1, weeks 6 and 12).
*P $<$ 0.05, **P $<$ 0.01, ***P $<$ 0.001, ****P $<$ 0.0001 for group by time interaction parameters of the linear models with repeated measures.
Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group; As, arsenic; InAs, inorganic arsenic; DMAs, dimethyl-arsenical species; MMAs, monomethyl-arsenical species.

person change in the proportion of urinary As metabolites between baseline and weeks 1, 6, and 12 differed between the 400FA, 800FA, and creatine+400FA treatment groups. Changes in the proportion of all metabolites from baseline to week 1 or weeks 6 and 12 did not differ between the 400FA and 800FA groups, or between the 400FA and creatine+400FA groups, with the exception that the decrease in %MMAs from baseline to weeks 6 and 12 was larger in the 800FA group than the 400FA group ($P = 0.034$).

Table 3.4: Summary statistics at week 24 and mean within-person changes in As metabolite proportions by FA supplement groups.[1]

	400FA			800FA		
	Week 24 Mean ± SD	**Mean change (95% CI) (wk 24 - wk 12)**	**Mean change (95% CI) (wk 24 - wk 0)**	**Week 24 Mean ± SD**	**Mean change (95% CI) (wk 24 - wk 12)**	**Mean change (95% CI) (wk 24 - wk 0)**
ln(%InAs)						
Continued suppl.	2.51 ± 0.37	-0.05 (-0.12, 0.03)	-0.09 (-0.17, -0.02)	2.49 ± 0.36	-0.01 (-0.09, 0.07)	-0.12 (-0.19, -0.04)
Discontinued suppl.	2.57 ± 0.35	0.06 (-0.03, 0.14)	0.01 (-0.08, 0.09)	2.54 ± 0.46	0.09 (0.00, 0.19)	-0.02 (-0.11, 0.08)
%MMAs						
Continued suppl.	11.76 ± 4.04	0.31 (-0.39, 1.02)	-1.63 (-2.36, -0.90)	12.04 4.16	1.02 (0.31, 1.74)	-1.27 (-2.01, -0.54)
Discontinued suppl.	12.27 ± 4.28	1.11 (0.46, 1.76)	-0.83 (-1.52, -0.14)	12.24 ± 5.00	1.82 (1.02, 2.63)	-0.47 (-1.31, 0.37)
%DMAs						
Continued suppl.	74.98 ± 7.2	0.47 (-0.82, 1.76)	2.81 (1.54, 4.08)	75.08 ± 7.00	-0.91 (-2.24, 0.41)	2.66 (1.31, 4.01)
Discontinued suppl.	73.8 ± 7.07	-2.03 (-3.43, -0.63)	0.31 (-1.08, 1.69)	73.88 ± 7.96	-3.42 (-4.85, -1.98)	0.16 (-1.27, 1.58)

1. Mean within-person changes and 95% CIs were derived from parameters of the linear models with repeated measures for effect of rebound.

Abbreviations: 400FA, 400 μg FA/day treatment group, 800FA; 800 μg FA/day treatment group; As, arsenic; FA, folic acid; InAs, inorganic arsenic; DMAs, dimethyl-arsenical species; MMAs, monomethyl-arsenical species; suppl., supplementation.

Treatment effects in the second phase

Proportions of urinary As metabolites after FA cessation in the 400FA and 800FA groups are shown in Table 3.4. At week 24, mean ln(%InAs) and %DMAs in both FA dose groups that discontinued FA supplementation were not significantly different than baseline levels. From baseline to week 24, there was a significantly greater mean decrease in %InAs and increase in %DMAs in both groups that continued FA supplementation compared to groups that discontinued supplementation (%InAs: time by FA continuation interaction $P = 0.028$; %DMAs: time by FA continuation interaction $P = 0.0005$). A similar pattern was observed in %MMAs with a significant difference at week 24 between groups that discontinued and continued FA supplementation (time by FA continuation $P = 0.048$); however, %MMAs at week 24 remained lower than baseline in the 400FA group that discontinued FA supplementation ($P = 0.018$).

3.6 Discussion

The purpose of this RCT was to investigate the effect of FA and creatine supplementation on As methylation among a population of mixed folate-deficient and -sufficient individuals. FA supplementation significantly increased As methylation; at weeks 6 and 12, there was a greater decrease in ln(%InAs) and %MMAs and increase in %DMAs among groups that had received 400 or 800 μg FA/day compared to placebo. The higher dose of FA had a greater effect on As methylation, as the decrease in %MMAs from baseline to weeks 6 and 12 was larger in the 800FA group than the 400FA group. Previously published findings from this trial demonstrated that supplementation of 800 μg FA/day, but not 400 μg FA/day, decreased total bAs concentrations to a significantly greater extent than placebo (Peters et al., 2015a).

These findings are in agreement with our previous 12-week RCT of the effects

of supplementation with 400 μg FA/day to folate-deficient adults in Bangladesh who did not receive As removal filters (Gamble et al., 2006). In that study, FA supplementation resulted in significant decreases in log(%InAs) and %MMAs, and increases in %DMAs in urine as compared to placebo. Arsenic methylation capacity was also assessed by the primary methylation index (PMI, calculated as MMAs/InAs) and secondary methylation index (SMI, calculated as DMAs/MMAs). Similar to the current study, effects were observed as early as one week after treatment began; SMI significantly increased by the week 1 follow-up in the group receiving FA compared to placebo.

Results from this trial provide evidence supporting the hypothesis that FA supplementation increases As methylation capacity, presumably by enhancing the synthesis of SAM, the methyl donor for As methylation, through provision of folate-derived methyl groups for one-carbon metabolism. Notably, effects of FA supplementation were observed in a population that was predominantly folate sufficient (plasma folate $\geq$ 9 nmol/L in 80.2% of total participants). Although the study was powered to detect differences in the full sample of both folate-deficient and -sufficient individuals, stratified analyses suggest that results were not driven by changes among participants who were folate-deficient at baseline. At week 12, the mean changes in percent As metabolites were similar between the folate-deficient and -sufficient strata in the groups receiving FA supplementation. In the 400FA treatment group, the mean changes in As metabolite proportions for participants with baseline deficient (N = 31) and sufficient folate status (N = 102) were, respectively: log(%InAs): -0.08, -0.04; %MMAs: -1.91 -1.98; %DMAs: 2.67, 2.25. Similarly, in the 800FA treatment group, the mean changes for folate deficient (N = 23) and sufficient participants (N = 106) were, respectively: log(%InAs): -0.07, -0.13; %MMAs: -1.96, -2.26; %DMAs: 2.83, 3.72.

The results suggest that percent urinary As metabolites returned to baseline levels 12 weeks after cessation of FA supplementation, as mean %InAs and %DMAs

were not significantly different than baseline levels in those groups that discontinued FA supplementation. However, as previously reported, although bAs concentrations did increase after discontinuation of FA supplementation, the increase did not achieve statistical significance (Peters et al., 2015a). In light of results regarding urinary As metabolites, the follow-up period may not have been long enough to observe a complete rebound in bAs after cessation of FA supplementation.

Creatine supplementation moderately influenced the proportion of As metabolites in urine. Although no significant differences were observed in the change in ln(%InAs) and %DMAs between the creatine and placebo groups at any time point, the decrease in %MMAs was significantly greater at week 1 and weeks 6 and 12 in the creatine group than the placebo group. In previously published results, creatine supplementation did not affect total bAs (Peters et al., 2015a). Our a priori hypothesis was that creatine supplementation would increase As methylation capacity by decreasing creatine biosynthesis, a major consumer of SAM methyl groups (Brosnan et al., 2011), and multiple studies have reported strong positive cross-sectional associations between urinary creatinine and As methylation capacity (Gamble et al., 2005, 2006; Ahsan et al., 2007; Hall et al., 2009; Basu et al., 2011; Peters et al., 2014). In the current study, we also observed significant correlations between urinary creatinine and %InAs and %DMAs. As we reported previously, supplementation of 3 g creatine/day did decrease creatine biosynthesis in this study population; concentrations of plasma guanidinoacetate, a precursor to creatine in the creatine biosynthetic pathway, significantly decreased with creatine supplementation, indicating a reduction in creatine biosynthesis (Peters et al., 2015b). The current findings lend support to the hypothesis that a decrease in creatine biosynthesis achieved through creatine supplementation may influence As methylation, however, the magnitude of the effects may be somewhat tempered by long-range allosteric regulation of one-carbon metabolism that regulate

intracellular SAM concentrations. For example, SAM inhibits methylenetetrahydrofolate reductase (MTHFR) (Jencks and Mathews, 1987), thereby decreasing synthesis of 5-MTHF and releasing 5-MTHF mediated inhibition of glycine N-methyltransferase (GNMT) activity (Wagner et al., 1985). GNMT serves as a major regulator of SAM concentrations through utilization of SAM for the nonessential conversion of glycine to sarcosine (Wagner et al., 1985; Luka et al., 2009). In an in silico experiment using a mathematical model of one-carbon metabolism, when creatine synthesis is set to zero, flux through GNMT increases by 30%. And, SAM concentrations are further modulated through increased flux through transsulfuration pathway because SAM stimulates cystathionine β-synthase (Reed et al., 2015). Additional unknown mechanisms may contribute to the strong cross-sectional associations between urinary creatinine and As methylation capacity. For example, it is possible that urinary creatinine is somehow related to renal tubular reabsorption of InAs, a topic that is relatively understudied in the scientific literature.

As noted above, this study was powered to test the effects of FA and creatine supplementation on a population of mixed folate-deficient and -sufficient individuals. Significant effects were observed among folate-sufficient participants receiving FA supplementation. In addition, the trial focused on the effects of increasing the recruitment of methyl groups to enhance one-carbon metabolism. However, As methylation capacity may also be altered by increasing the availability of other methyl donors involved in one-carbon metabolism (Ueland, 2011). Further research is needed to explore the impact of co-supplementation of FA with other methyl donors, such as choline and betaine, and other micronutrients related to one-carbon metabolism, such as vitamins B_{12}, B_2, and B_6. Furthermore, treatment effects may differ by baseline nutritional status of these micronutrients.

The study limitations include poor compliance with water filter use over

time and the possibility that some urine samples were diluted with water. However, it is unlikely that either of these factors differed between treatment groups due to randomization, and therefore they are not expected to influence the studys conclusions.

3.7 Conclusion

The importance of nutritional factors in modifying As metabolism and toxicity has emerged over the past 20 years (Abernathy et al., 1999), as well as the potential use of nutritional supplements to mitigate the health effects of chronic As exposure (Kile and Ronnenberg, 2008). Remediation of As exposure is the most important approach to limiting health effects (World Health Organization, 2012), however, removal of As from drinking water may not be immediately feasible in all real-world situations and exposure to unsafe concentrations of As in drinking water persists in many regions of the world (Naujokas et al., 2013). Furthermore, adverse health outcomes associated with chronic As exposure persist after exposure has ended (Steinmaus et al., 2013, 2014). Additional approaches to lower bAs, increase As methylation and potentially reduce As toxicity are needed. In this trial, FA supplementation significantly increased As methylation capacity, resulting in an increased proportion of urinary DMAs, the least toxic and most readily excreted As metabolite. A significant change in the proportion of InAs and DMAs was not observed in the creatine group, although findings suggest that creatine supplementation decreased %MMAs to a lesser extent than FA. While these findings are informative, further research is needed to fully understand the strong cross-sectional relationships previously observed between urinary creatinine and arsenic methylation (Gamble et al., 2005; Ahsan et al., 2007; Hall et al., 2009; Peters et al., 2014).

Nutritional interventions may conceivably reduce long-term health risks of As exposure. For example, folate deficiency has been associated with the subsequent devel-

opment of skin lesions (Pilsner et al., 2009). FA fortification of grains has been shown to dramatically decrease the prevalence of folate deficiency (Odewole et al., 2013; Barnabé et al., 2015) and has been mandated in 87 countries (Zimmerman and Lu, 2015), however Bangladesh is not among them. Countries with endemic As exposure, such as Bangladesh, may benefit from FA fortification to reduce folate deficiency in general, and as a cost-effective method of partially reducing As toxicity as one component of comprehensive As mitigation programs.

3.8 Bibliography

Abernathy, C O; Liu, Y P; Longfellow, D; Aposhian, H V; Beck, S; Fowler, B; Goyer, R; Menzerl, R; Rossman, T; Thompson, C, and Waalkes, M. Arsenic: health effects, mechanisms of actions, and research issues. *Environmental Health Perspectives*, 107: 593–597, 1999.

Ahsan, H; Chen, Y; Parvez, F; Argos, M; Hussain, A I; Momotaj, H; Levy, D; van Geen, A; Howe, G, and Graziano, J. Health Effects of Arsenic Longitudinal Study (HEALS): Description of a multidisciplinary epidemiologic investigation. *Journal of Exposure Science and Environmental Epidemiology*, 16(2):191–205, 2006. doi: 10.1038/sj.jea.7500449.

Ahsan, H; Chen, Y; Kibriya, M G; Slavkovich, V; Parvez, F; Jasmine, F; Gamble, M V, and Graziano, J H. Arsenic metabolism, genetic susceptibility, and risk of premalignant skin lesions in Bangladesh. *Cancer Epidemiology Biomarkers and Prevention*, 16(6):1270–1278, 2007. doi: 10.1158/1055-9965.EPI-06-0676.

Barnabé, A; Aléssio, A; Bittar, L; de Moraes Mazetto, B; Bicudo, A M; de Paula, E V; Höehr, N, and M, Annichino-Bizzacchi J. Folate, vitamin B12 and homocysteine status in the post-folic acid fortification era in different subgroups of the Brazilian population attended to at a public health care center. *Nutrition Journal*, 14(1):19, 2015. doi: 10.1186/s12937-015-0006-3.

Basu, A; Mitra, S; Chung, J; Guha Mazumder, DN; Ghosh, N; Kalman, D; von Ehrenstein, O S; Steinmaus, C; Liaw, J, and Smith, A H. Creatinine, diet, micronutrients, and arsenic methylation in West Bengal, India. *Environmental Health Perspectives*, 119(9):1308–1313, 2011. doi: 10.1289/ehp.1003393.

Brosnan, J T; da Silva, R P, and Brosnan, M E. The metabolic burden of creatine synthesis. *Amino Acids*, 40(5):1325–31, 2011. doi: 10.1007/s00726-011-0853-y.

Challenger, F. Biological methylation. *Chemical Reviews*, 36(3):315–361, 1945. doi: 10.1021/cr60115a003.

Gamble, M V; Liu, X; Ahsan, H; Pilsner, R; Ilievski, V; Slavkovich, V; Parvez, F; Levy, D; Factor-Litvak, P, and Graziano, J H. Folate, homocysteine, and arsenic metabolism in arsenic-exposed individuals in Bangladesh. *Environmental Health Perspectives*, 113(12):1683–8, 2005. doi: 10.1289/ehp.8084.

Gamble, M V; Liu, X; Ahsan, H; Pilsner, J R; Ilievski, V; Slavkovich, V; Parvez, F; Chen, Y; Levy, D; Factor-Litvak, P, and Graziano, J H. Folate and arsenic metabolism: a double-blind, placebo-controlled folic acidsupplementation trial in Bangladesh. *American Journal of Clinical Nutrition*, 84(5):1093–101, 2006. doi: 10.1093/ajcn/84.5.1093.

Gamble, M V; Liu, X; Slavkovich, V; Pilsner, J R; Ilievski, V; Factor-Litvak, P; Levy, D; Alam, S; Islam, M; Parvez, F; Ahsan, H, and Graziano, J H. Folic acid supplementation lowers blood arsenic. *American Journal of Clinical Nutrition*, 86(4):1202–9, 2007. doi: 10.1093/ajcn/86.4.1202.

Government of the People's Republic of Bangladesh; Department for International Development, and British Geological Survey. Arsenic contamination of groundwater in Bangladesh, Vol. 1: Summary, 2001.

Hall, M N and Gamble, M V. Nutritional manipulation of one-carbon metabolism: effects on arsenic methylation and toxicity. *Journal of Toxicology*, (11):1–11, 2012. doi: 10.1155/2012/595307.

Hall, M N; Liu, X; Slavkovich, V; Ilievski, V; Pilsner, J R; Alam, S; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate, cobalamin, cysteine, homocysteine, and arsenic metabolism among children in Bangladesh. *Environmental Health Perspectives*, 117(5), 2009. doi: 10.1289/ehp.0800164.

Jencks, D A and Mathews, R G. Allosteric inhibition of methylenetetrahydrofolate reductase by adenosylmethionine. Effects of adenosylmethionine and NADPH on the equilibrium between active and inactive forms of the enzyme and on the kinetics of approach to equilibrium. *Journal of Biological Chemistry*, 262(6):2485–93, 1987.

Kile, M L and Ronnenberg, A G. Can folate intake reduce arsenic toxicity? *Nutrition Reviews*, 66(6):349–53, 2008. doi: 10.1111/j.1753-4887.2008.00043.x.

Loffredo, C A; Aposhian, H V V; Cebrian, M E; Yamauchi, H, and Silbergeld, E K. Variability in human metabolism of arsenic. *Environmental Research*, 92(2):85–91, 2003. doi: 10.1016/S0013-9351(02)00081-6.

Luka, Z; Mudd, S H, and Wagner, C. Glycine N-methyltransferase and regulation of S-adenosylmethionine levels. *Journal of Biological Chemistry*, 284(34):22507–22511, 2009. doi: 10.1074/jbc.R109.019273.

Miller, R C; Brindle, E; Holman, D J; Shofer, J; Klein, N A; Soules, M R, and O 'connor, K A. Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. *Clinical Chemistry*, 50(5):924–932, 2004. doi: 10.1373/clinchem.2004.032292.

Moe, B; Peng, H; Lu, X; Chen, B; Chen, L WL; Gabos, S; Li, X-F, and Le, X C. Comparative cytotoxicity of fourteen trivalent and pentavalent arsenic species determined using real-time cell sensing. *Journal of Environmental Sciences*, 49:113–124, 2016. doi: 10.1016/j.jes.2016.10.004.

Naujokas, M F.; Anderson, B; Ahsan, H; Vasken Aposhian, H; Graziano, J H.; Thompson, C, and Suk, W A. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environmental Health Perspectives*, 121(3):295–302, 2013. doi: 10.1289/ehp.1205875.

Odewole, O A; Williamson, R S; Zakai, N A; Berry, R J; Judd, S E; Qi, Y P; Adedinsewo, D A, and Oakley, G P. Near-elimination of folate-deficiency anemia by mandatory folic acid fortification in older US adults: reasons for geographic and racial differences in stroke study 2003-2007. *American Journal of Clinical Nutrition*, 98(4): 1042–7, 2013. doi: 10.3945/ajcn.113.059683.

Peters, B A.; Hall, M N.; Liu, X; Neugut, Y D; Pilsner, J R; Levy, D; Ilievski, V; Slavkovich, V; Islam, T; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Creatinine, arsenic metabolism, and renal function in an arsenic-exposed population in Bangladesh. *PLoS ONE*, 9(12):e113760, 2014. doi: 10.1371/journal.pone.0113760.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Sanchez, T; van Geen, A; Mey, J L; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T; Balac, O; Ilievski, V; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folic acid and creatine as therapeutic approaches to lower blood arsenic: a randomized controlled trial. *Environmental Health Perspectives*, 123(12):1294–1301, 2015a. doi: 10.1289/ehp.1409396.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M Nasir; Islam, T; Ilievski, V; Graziano, J H, and Gamble, M V. Low-dose creatine supplementation lowers plasma guanidinoacetate, but not plasma homocysteine, in a double-blind, randomized, placebo-controlled trial. *Journal of Nutrition*, 145(10): 2245–52, 2015b. doi: 10.3945/jn.115.216739.

Petrick, J S; Ayala-Fierro, F; Cullen, W R; Carter, D E, and Vasken Aposhian, H. Monomethylarsonous acid (MMA(III)) is more toxic than arsenite in Chang human hepatocytes. *Toxicology and Applied Pharmacology*, 163(2):203–207, 2000. doi: 10.1006/taap.1999.8872.

Pierce, B L; Tong, L; Argos, M; Gao, J; Jasmine, F; Roy, S; Paul-Brutus, R; Rahaman, R; Rakibuz-Zaman, M; Parvez, F; Ahmed, A; Quasem, I; Hore, S K; Alam, S; Islam, T; Harjes, J; Sarwar, G; Slavkovich, V; Gamble, M V; Chen, Y; Yunus, M; Rahman, M; Baron, J A; Graziano, J H, and Ahsan, H. Arsenic metabolism efficiency has a causal role in arsenic toxicity: Mendelian randomization and gene-environment interaction. *International Journal of Epidemiology*, 42(6):1862–1872, 2013. doi: 10.1093/ije/dyt182.

Pilsner, J R; Liu, X; Ahsan, H; Ilievski, V; Slavkovich, V; Levy, D; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. *Environmental Health Perspectives*, 117(2):254–60, 2009. doi: 10.1289/ehp.11872.

R Core Team. R: A Language and Environment for Statistical Computing. Technical report, R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `https://www.r-project.org/`.

Reed, M C; Gamble, M V; Hall, M N, and Nijhout, H F. Mathematical analysis of the regulation of competing methyltransferases. *BMC Systems Biology*, 9, 2015. doi: 10.1186/s12918-015-0215-6.

Reuter, W; Davidowski, L, and Neubauer, K. Speciation of five arsenic compounds in urine by HPLC/ICP-MS, PerkinElmerSCIEX. Technical report, PerkinElmer Life and Analytical Sciences, 2003.

Sanchez, T R; Levy, D; Shahriar, M H; Uddin, M N; Siddique, A B; Graziano, J H; Lomax-Luu, A; van Geen, A, and Gamble, M V. Provision of well-water treatment units to 600 households in Bangladesh: a longitudinal analysis of urinary arsenic indicates fading utility. *Science of The Total Environment*, 563:131–137, 2016. doi: 10.1016/j.scitotenv.2016.04.112.

Steinmaus, C; Yuan, Y; Kalman, D; Rey, O A; Skibola, C F; Dauphine, D; Basu, A; Porter, K E; Hubbard, A; Bates, M N; Smith, M T, and Smith, A H. Individual differences in arsenic metabolism and lung cancer in a case-control study in Cordoba, Argentina. *Toxicology and Applied Pharmacology*, 247(2):138–45, 2010. doi: 10.1016/j.taap.2010.06.006.

Steinmaus, C; Ferreccio, C; Acevedo, J; Yuan, Y; Liaw, J; Duran, V; Cuevas, S; Garcia, J; Meza, R; Valdes, R; Valdes, G; Benitez, H; VanderLinde, V; Villagra, V; Cantor, K P; Moore, L E; Perez, S G; Steinmaus, S, and Smith, A H. Increased lung and bladder cancer incidence in adults after *in utero* and early-life arsenic exposure. *Cancer Epidemiology Biomarkers and Prevention*, 23(8):1529–38, 2014. doi: 10.1158/1055-9965.EPI-14-0059.

Steinmaus, C M; Ferreccio, C; Romo, J A; Yuan, Y; Cortes, S; Marshall, G; Moore, L E; Balmes, J R; Liaw, J; Golden, T, and Smith, A H. Drinking water arsenic in northern Chile: high cancer risks 40 years after exposure cessation. *Cancer Epidemiology Biomarkers and Prevention*, 22(4):623–30, 2013. doi: 10.1158/1055-9965.EPI-12-1190.

Syracuse Research Corporation. Toxicological Profile for Arsenic (Update). Technical report, U.S. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances Disease Registry, Atlanta, GA, 2007.

Thomas, D J; Li, J; Waters, S B; Xing, W; Adair, B M; Drobna, Z; Devesa, V, and Styblo, M. Arsenic (+3 oxidation state) methyltransferase and the methylation of arsenicals. *Experimental biology and medicine (Maywood, N.J.)*, 232(1):3–13, 2007.

Ueland, P M. Choline and betaine in health and disease. *Journal of Inherited Metabolic Disease*, 34(1):3–15, 2011. doi: 10.1007/s10545-010-9088-4.

Vahter, M. Methylation of inorganic arsenic in different mammalian species and population groups. *Science progress*, 82:69–88, 1999.

Wagner, C; Briggs, W T, and Cook, R J. Inhibition of glycine N-methyltransferase activity by folate derivatives: implications for regulation of methyl group metabolism. *Biochemical and Biophysical Research Communications*, 127(3):746–52, 1985.

World Health Organization. Arsenic, 2012. URL `http://www.who.int/mediacentre/factsheets/fs372/en/`.

Yu, R C; Hsu, K-H; Chen, C-J, and Froines, J R. Arsenic methylation capacity and skin cancer. *Cancer Epidemiology, Biomarkers and Prevention*, 9:1259–1262, 2000.

Zhang, Q; Li, Y; Liu, J; Wang, D; Zheng, Q, and Sun, G. Differences of urinary arsenic metabolites and methylation capacity between individuals with and without skin lesions in Inner Mongolia, Northern China. *International Journal of Environmental Research and Public Health*, 11(7):7319–32, 2014. doi: 10.3390/ijerph110707319.

Zimmerman, S and Lu, C. Defeating Anemia: 2015 Year in Review. Technical report, Food Fortification Initiative, Atlanta, GA, 2015.

3.9 Appendix

CONSORT
TRANSPARENT REPORTING of TRIALS

Supplemental Figure 1: CONSORT Flow Diagram

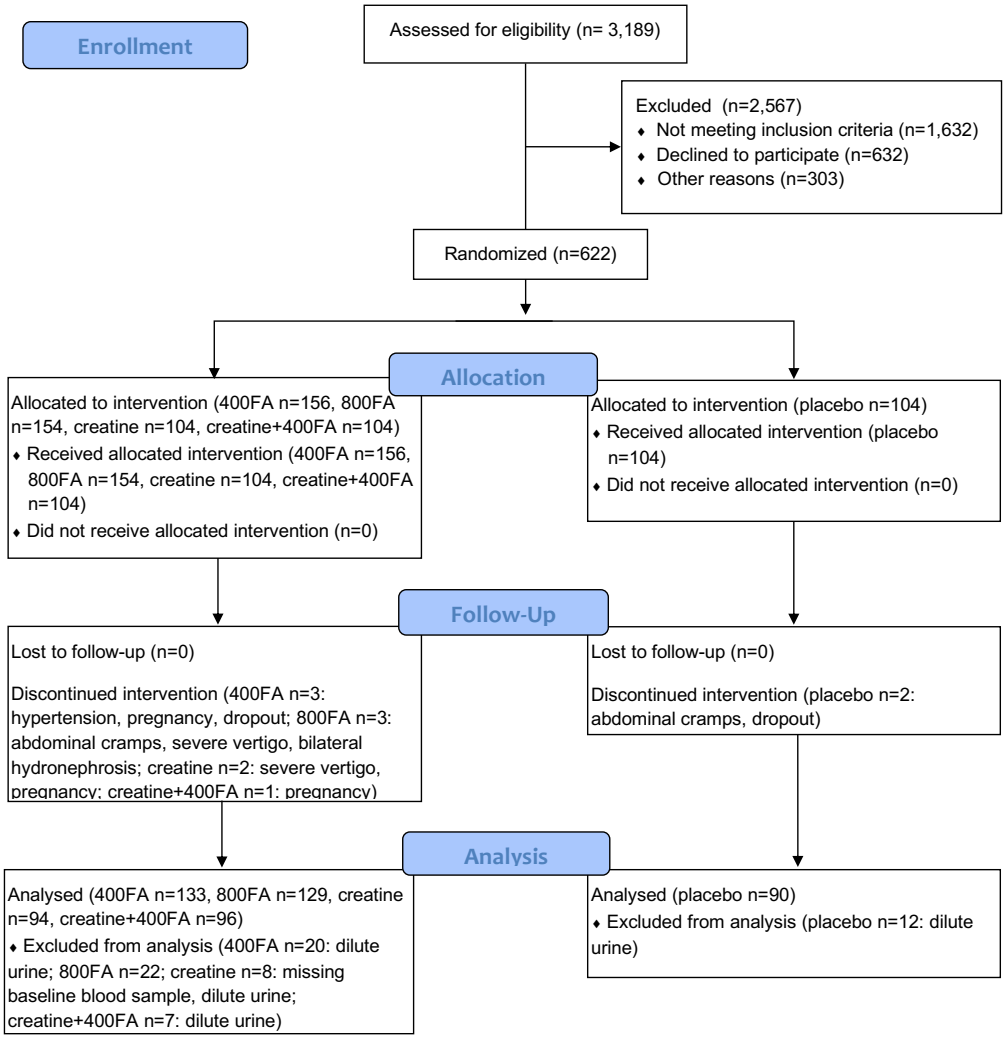


Figure 3.5: CONSORT flowchart for FACT.

Chapter 4

Betaine and choline status modify the effect of folic acid and creatine supplementation on arsenic methylation in a randomized controlled trial of Bangladeshi adults

Anne K. Bozack,[1], Caitlin G. Howe,[2] Megan N. Hall,[4] Xinhua Liu,[4] Vesna Slavkovich,[1], Vesna Ilievski,[1] Angela M. Lomax-Luu,[1] Faruque Parvez,[1] Abu B. Siddique,[5] Hasan Shahriar,[5] Mohammad N. Uddin,[5] Tariqul Islam,[5] Joseph H. Graziano,[1] Mary V. Gamble[1]

Affiliations: 1. Dept. of Environmental Health Sciences, Columbia University; 2. Dept. of Preventive Medicine, University of Southern California; 3. Department of Epidemiology, Columbia University, and Departments of Environmental and

Occupational Health Sciences & Epidemiology, SUNY Downstate School of Public Health; 4. Department of Biostatistics, Columbia University; 5. Columbia University Arsenic Project in Bangladesh, Dhaka, Bangladesh.

Submitted to the *Journal of Nutrition*.

4.1 Abstract

Background: Arsenic (As) is a human toxicant and carcinogen. Methylation of inorganic As (InAs) to monomethyl- (MMAs) and dimethyl-arsenical species (DMAs) facilitates urinary As elimination. Folate and creatine, one-carbon metabolism (OCM) nutrients, influenced As methylation in a randomized controlled trial (RCT): folic acid (FA) lowered blood As (bAs) and increased As methylation, and creatine decreased urinary %MMAs. Choline and betaine are alternative methyl donors in OCM.

Aim: This study examined if baseline plasma folate, choline, betaine, and vitamin B_{12} status modify the effects of 12 weeks of FA and creatine supplementation on changes in homocysteine, guanidinoacetate (GAA), total bAs, and urinary As metabolite proportions and indices.

Methods: In a RCT, 622 participants were assigned to receive 400 or 800 μg FA, 3 g creatine, 400 μg FA + 3 g creatine, or placebo daily. All participants received As-removal water filters.

Results: Over 12 weeks, relative to placebo, 400 and 800 μg FA/day were associated with greater mean increases in %DMAs among participants with baseline betaine concentrations below the median than those with levels above the median ($FDR < 0.05$). 400 μg FA/day was associated with a greater decrease in homocysteine among participants with plasma folate concentrations below, compared with those above, the median ($FDR < 0.03$). Creatine treatment was associated with a significant decrease in %MMAs compared to placebo among participants with choline concentrations below the median ($P = 0.04$), but not among participants above the median ($P = 0.94$); this effect did not significantly differ between strata ($P = 0.10$).

Conclusions: Effects of FA and creatine supplementation on As methylation capacity were greater among individuals with low betaine and choline status, respectively. These findings reveal that the efficacy of nutritional interventions with FA and

creatine to facilitate As methylation is modified by choline and betaine nutritional status.

Acknowledgments

We thank our staff, the fieldworkers, and the study participants in Bangladesh, without whom this work would not have been possible.

4.2 Introduction

Chronic arsenic (As) exposure through drinking water is a major public health concern. Over 140 million people in more than 70 countries are exposed to As concentrations > 10 μg/L, the World Health Organization (WHO) guideline (Naujokas et al., 2013; World Health Organization, 2012). In Bangladesh, approximately 40 million people are exposed to drinking water As concentrations exceeding 10 μg/L (Bangladesh Bureau of Statistics and United Nation Children's Fund, 2015). Chronic As exposure has been associated with increased risk of numerous health conditions including cardiovascular disease, diabetes, skin lesions (melanosis, leukomelanosis, and keratosis), cancers (bladder, kidney, liver, lung, skin, and prostate), and impaired intellectual function (Benbrahim-Tallaa and Waalkes, 2008; Naujokas et al., 2013).

Methylation of inorganic As (InAs) to mono- and di-methyl arsenical species facilitates urinary As excretion (Tice et al., 1997; Vahter and Marafante, 1987). Ingested InAs is metabolized through a series of reduction and oxidative methylation reactions (Challenger, 1945). $InAs^{III}$ is methylated to monomethylarsonic acid ($MMAs^{V}$), reduced to $MMAs^{III}$, and methylated to dimethylarsinic acid ($DMAs^{V}$). These sequential reactions are catalyzed by arsenic-3-methyltransferase (AS3MT) using the methyl donor *S*-adenosylmethionine (SAM) (Thomas et al., 2004) (Figure 4.1). Toxicological studies have demonstrated that $MMAs^{III}$ is the most cytotoxic and genotoxic As species (Moe et al., 2016; Petrick et al., 2000). Due to rapid oxidation, it is difficult to distinguish between $MMAs^{III}$ and $MMAs^{V}$ in human studies. However, in epidemiological studies, a higher proportion of $MMAs^{III+V}$ (%MMAs) and lower %DMAs in urine has been associated with increased risks for bladder, breast, lung, and skin cancers; skin lesions; peripheral vascular disease; and atherosclerosis (Kuo et al., 2017; Steinmaus et al., 2010).

Arsenic metabolism efficiency varies between individuals and is influenced by

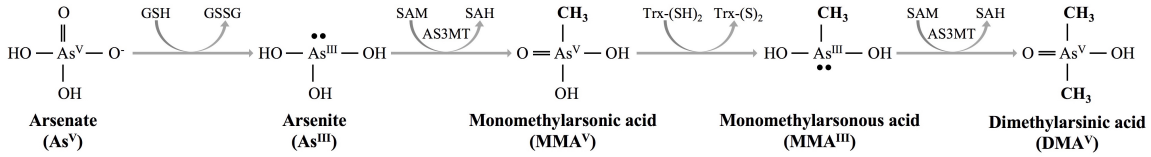


Figure 4.1: Arsenic methylation. As^{III} is methylated to form $MMAs_V$ by AS3MT using the methyl donor SAM. $MMAs^V$ is subsequently reduced to $MMAs^{III}$ and methylated to form $DMAs^V$. Abbreviations: As^{III}, arsenite; AS3MT, arsenic arsenic (+3 oxidation state) methyltransferase; $DMAs^V$, dimethylarsinic acid; InAs, inorganic As; $MMAs^V$, monomethylarsonic acid; $MMAs^{III}$, monomethylarsonous acid; SAM, *S*-adenosylmethionine.

one-carbon metabolism (OCM), the biochemical pathway that synthesizes SAM. Recruitment of one-carbon units into OCM is influenced by folate; OCM is also influenced by micronutrients that act as cofactors (e.g., vitamin B_{12}) or alternative methyl donors (choline and betaine). A one-carbon unit is transferred from 5-methyl-tetrahydrofolate (5-methyl-THF), the most prevalent form of naturally occurring folate, to homocysteine by methionine synthase (MTR) using the cofactor vitamin B_{12} to form methionine (Figure 4.2). Methionine is subsequently activated to SAM. In the liver, betaine can serve as an alternative methyl donor to remethylate homocysteine. Betaine is obtained through diet or synthesized endogenously from choline, which in turn can be obtained from food or synthesized from phosphatidylcholine (PC). When folate status is low, the use of betaine for homocysteine remethylation is increased (Niculescu and Zeisel, 2002).

Dietary creatine may also influence the availability of SAM. An estimated 50% of SAM is consumed by creatine biosynthesis, in which guanidinoacetate (GAA) methyltransferase (GAMT) (Brosnan, da Silva, and Brosnan 2011) catalyzes the methylation of GAA to form creatine and *S*-adenosylhomocysteine (SAH) (Figure 4.2). Dietary creatine reduces GAA biosynthesis through the pre-translational inhibition of arginine:glycine amidinotransferase (AGAT), which catalyzes the synthesis of GAA from arginine and glycine (McGuire et al., 1984). Dietary sources, predominantly meat,

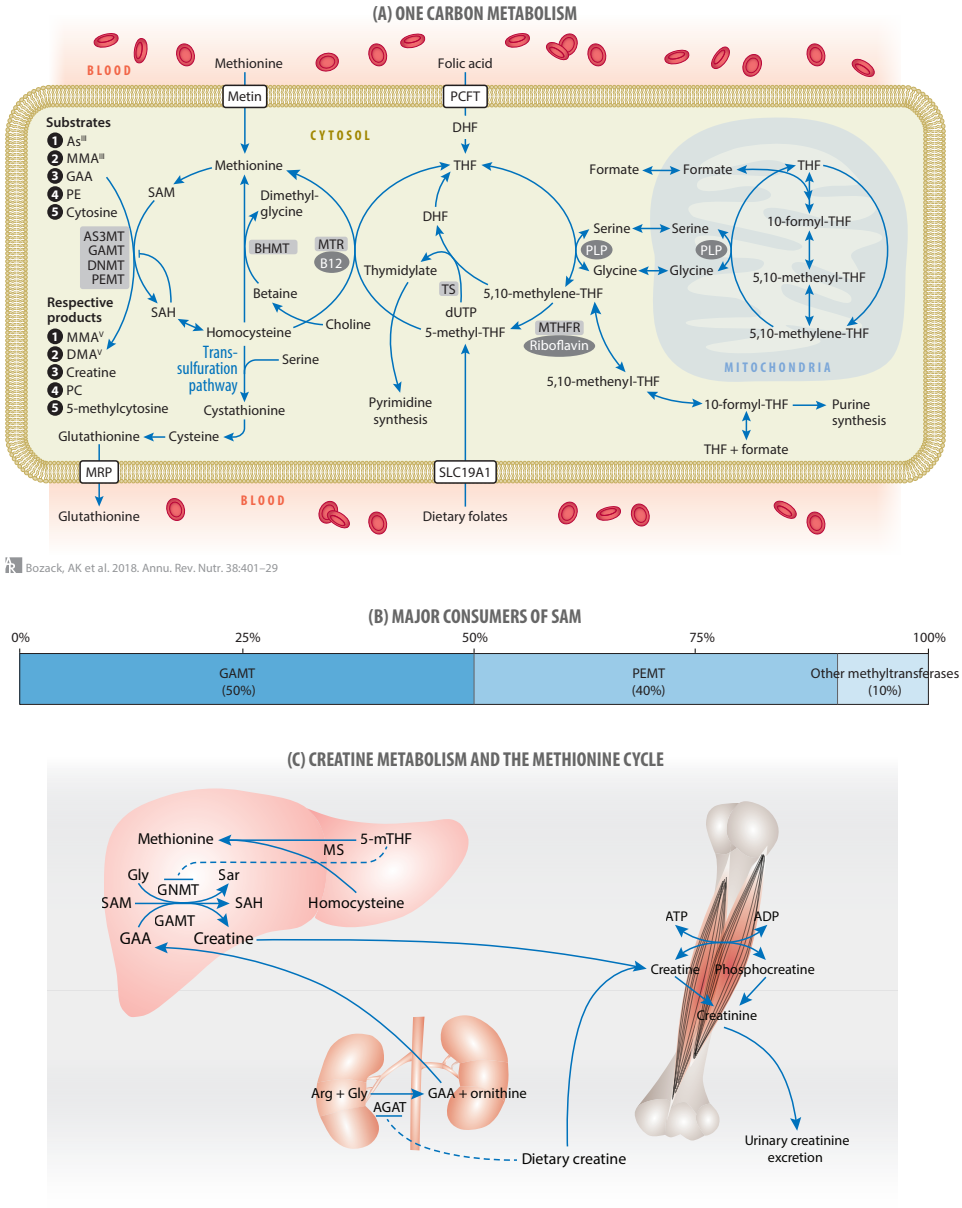


Figure 4.2: (Caption next page.)

Figure 4.2: (Previous page.) **(A) One-carbon metabolism (OCM).** FA is reduced to DHF and THF by dihydrofolate reductase. 5,10-methylene-THF is formed by serine hydroxymethyl-transferase through the transfer of one-carbon units from serine to THF, which is for thymidylate synthesis or reduced to 5-mTHF. Folate obtained through the diet can enter one-carbon metabolism as 5-mTHF. A one-carbon unit is transferred from 5-mTHF to homocysteine by MTR using vitamin B_{12} as a cofactor to form methionine and THF. Homocysteine can also be remethylated in the liver by BHMT using betaine as the methyl donor. Methionine is activated to from SAM by methionine adenosyltransferase enzymes. SAM serves as the methyl donor for numerous reactions including arsenic methylation, the biosynthesis of creatine from GAA, and DNA methylation, generating the methylated products and SAH. SAH, which serves as a product inhibitor for most methyltransferase enzymes, hydrolyzed to homocysteine, and can either be remethylated to methionine or be directed towards the transsulfuration pathway. Adapted with permission from (16). **(B) Major consumers of SAM.** An estimated 50% of SAM is consumed by the final step of endogenous creatine synthesis by GAMT, and 40% of SAM is consumed by phosphatidylcholine biosynthesis by PEMT. **(C) Creatine metabolism and the methionine cycle.** In the kidney, AGAT transfers an amidino group from arginine to glycine, producing GAA and ornithine. Dietary and/or supplemental creatine reduces GAA biosynthesis through the pretranslational inhibition of AGAT. GAA is released from the kidney and taken up by the liver where it is methylated by GAMT using the methyl donor SAM to form creatine and SAH. SAH is hydrolyzed to homocysteine. 5-mTHF can regulate SAM and SAH levels through potent inhibition of GNMT. Creatine is transported to tissues including skeletal muscle, heart, and brain, and phosphorylated to phosphocreatine. Creatine and phosphocreatine are converted to creatinine through a nonenzymatic reaction and excreted in urine. Abbreviations: 5-mTHF, 5-methyl-tetrahydrofolate; AGAT, arginine:glycine amidinotransferase; BHMT, betaine homocysteine methyltransferase; DHF, dihydrofolate; GAA, guanidinoacetate; GAMT, guanidinoacetate N-methyltransferase; MTR, methionine synthase; OCM, one-carbon metabolism; PEMT, phosphatidylethanolamine N-Methyltransferase; SAH, *S*-adenosylhomocysteine; SAM, *S*-adenosylmethionine; THF, tetrahydrofolate.

provide approximately half of the daily requirement for creatine, but this proportion is lower for vegetarians (Brosnan et al., 2011).

Our group and others have reported that dietary folate intake and folate status are positively associated with As methylation capacity (reviewed in Bozack et al. 2018b). We have also studied the effect of FA supplementation on As metabolism and elimination in Bangladeshi adults. In a 12-week randomized controlled trial (RCT) among participants with low plasma folate (< 9 nmol/L), 400 μg FA/day supplemen-

tation, the U.S. recommended daily allowance (RDA), was associated with a larger increase in urinary %DMAs and decreases in ln(%InAs), %MMAs (Gamble et al., 2006), total blood As (bAs) concentration, and blood MMAs concentration compared to placebo (Gamble et al., 2007). In the Folic Acid and Creatine Trial (FACT), an RCT among adults recruited independent of folate status, we observed a larger increase in urinary %DMAs and decreases in %InAs and %MMAs after 12 weeks of 400 or 800 μg FA/day supplementation (Bozack et al., 2018a), and a larger decrease in ln(bAs) with 800 μg FA/day supplementation compared to placebo (Peters et al., 2015a). Supplementation with 400 and 800 μg FA resulted in significant increases in plasma betaine, illustrating the sparing effect of FA on betaine for homocysteine remethylation (Hall et al., 2016).

The associations between additional OCM-related micronutrients and As methylation capacity have been investigated in human studies (Bozack et al., 2018b). Urinary creatinine, a product of creatine metabolism and a biomarker of dietary creatine intake and endogenous creatine biosynthesis, has been consistently associated with lower %InAs and higher %DMAs in urine in cross-sectional analyses (Basu et al., 2011; Bozack et al., 2018a; Gamble et al., 2005, 2006; Hall et al., 2009, 2006; Kile et al., 2009; Pilsner et al., 2009). In the FACT study, 3 g/day creatine supplementation was associated with a larger decrease in plasma GAA compared to placebo, indicating downregulation of endogenous creatine synthesis (Peters et al., 2015b). Creatine supplementation was associated with a larger decrease in urinary %MMAs compared to placebo at 6 and 12 weeks, but, surprisingly, was not associated with changes in %InAs or %DMAs (Bozack et al., 2018a). The association between choline and betaine and As methylation capacity has been investigated using food frequency questionnaire data. Dietary choline, but not betaine, has been positively associated with As methylation capacity as measured by %InAs, %DMAs, DMAs/InAs (López-Carrillo et al.,

2016) and DMAs/MMAs (Heck et al., 2007; López-Carrillo et al., 2016) in urine. Findings regarding the association between vitamin B_{12} and the proportion of urinary As metabolites are less consistent; results differ in the direction and significance of the associations across studies (Hall et al., 2009; López-Carrillo et al., 2016; Spratlen et al., 2017).

It is not known if FA and creatine supplementation treatment effects are modified by OCM-related micronutrients. Given the reciprocal use of folate vs. choline/betaine for the remethylation of homocysteine, we hypothesized that participants with low baseline levels of OCM-related micronutrients would experience greater treatment effects due to a limited supply of methyl donors prior to treatment. The objectives of the analyses presented here are to determine if baseline folate, choline, betaine, and vitamin B_{12} status modify the effects of FA and creatine supplementation on changes in homocysteine, GAA, bAs concentration, and urinary As metabolite proportions and methylation indices.

4.3 Methods

Subjects

FACT was a randomized, double-blind, placebo-controlled trial to investigate the effects of FA and creatine supplementation on change in total bAs, and has been described in detail previously by Peters et al. (Peters et al., 2015a). Participants were randomly recruited from the Health Effects of Arsenic Longitudinal Study (HEALS) (Ahsan et al., 2006), a cohort of over 30,000 adults in Araihazar, Bangladesh. Participants were eligible for FACT if they were drinking from a household well with As concentration ≥ 50 μg/L for $\geq$ one year prior to enrollment. Participants were excluded if they were pregnant, taking nutritional supplements, or had proteinuria, renal disease,

diabetes, gastrointestinal problems, or other health issues.

Study design

A total of 622 participants were recruited. Participants were provided with READ-F As removal filters (READ-F filter; Brota Services International, Bangladesh) and were encouraged to use the filters for all drinking and cooking water during the trial (Sanchez et al., 2016). As previously described (Peters et al., 2015a), participants were assigned to one of five treatment groups: 400 μg FA/day (referred to hereafter as 400FA; N = 156), 800 μg FA/day (800FA; N = 154), 3 g creatine/day (creatine; N = 104), 3 g creatine and 400 μg FA/day (creatine+400FA; N = 104), and placebo (N = 104). The FA doses of 400 and 800 μg/day were selected to meet and exceed the U.S. RDA; the creatine dose of 3 g/day was selected to exceed daily creatine loss (approximately 2 g for 70 kg 20-39 year-old males) (Brosnan et al., 2011), to be sufficient to downregulate endogenous creatine synthesis. Supplements were provided by Atrium Innovations, Inc. (Westmount, Quebec).

During the first 12-week phase, participants received daily supplements or a placebo; during the second 12-week phase, participants in the FA treatment groups were randomly assigned to continue their FA treatment (400FA: N = 77; 800FA: N = 77) or to receive a placebo (400FA/placebo: N = 76; 800FA/placebo: N = 74), and participants in the creatine and creatine+400FA groups received a placebo to maintain the study blind.

Results regarding changes in bAs (Peters et al., 2015a) and urinary As methylation (Bozack et al., 2018a) have previously been published. The current analyses utilized data from the first phase of the trial to investigate the *a priori* hypothesis that baseline nutritional status modifies the association between FA and/or creatine supplementation and changes in total bAs, urinary As metabolite proportions and indices,

homocysteine, and GAA between baseline and week 12.

Ethics

The Columbia University Medical Center Institutional Review Board and the Bangladesh Medical Research Council approved the study protocol. Informed consent was obtained by staff physicians in Bangladesh.

Field work and participant follow-up

Field work was conducted in 2010-2012. Five pairs of field staff (an interviewer and physician) conducted recruitment and home visits during which venous blood (baseline and weeks 12 and 24) or urine samples (baseline and weeks 1, 6, 12, 13, 18, and 24) were collected. During daily home visits, health workers observed or inquired about participants taking the pills. Pill counts were conducted at weeks 12 and 24. Compliance was high (range: 79.1-100%; median: 99.5%; interquartile range: 98.3-100.0%) and did not differ substantially between treatment groups (Peters et al., 2015a).

Laboratory measures

Sample handling procedures and laboratory methods have previously been described in detail (Howe et al., 2017; Peters et al., 2015a,b). Venous blood samples were collected in EDTA vacutainer tubes, stored at 4°C in IsoRack cool packs (Brinkmann Instruments; Riverview, FL). Urine samples were collected in 50-mL acid-washed polypropylene tubes and stored in portable coolers. Samples were transported to our Araihazar field clinic within 4 hours. Blood plasma was separated using centrifugation. Blood and urine samples were shipped to Columbia University on dry ice and stored at -80°C and -20°C, respectively.

Total bAs was measured using Inductively Coupled Plasma Mass Spectrometry (ICP-MS) (PerkinElmer Elan DRC II; Waltham, MA; with an AS 93+ autosampler) (intra- and inter-assay CVs: 2.7% and 5.7%, respectively) (Hall et al., 2006). Plasma folate and vitamin B_{12} were measured by radioimmunoassay (SimulTRAC-SNB, MP Biomedicals) (intra- and inter-assay CVs: 5% and 13% for plasma folate; 6% and 17% for vitamin B_{12}).

High performance liquid chromatography (HPLC) with fluorescence detection was used to measure total plasma homocysteine (Pfeiffer et al., 1999) (intra- and inter-assay CVs: 5% and 7%) and plasma GAA (Carducci et al., 2002) (intra- and inter-assay CVs: 8% and 9%). Plasma choline and betaine were measured using liquid chromatographytandem mass spectrometry (LC-MS/MS) (Holm et al., 2003; Yan et al., 2012) (intra- and inter-assay CVs: 2.2% and 5.8% for plasma choline; 2.5% and 5.6% for plasma betaine).

Urinary arsenobetaine, arsenocholine, As^{III}, As^{V}, MMAs, and DMAs were separated by HPLC and measured by ICP-MS with dynamic reaction cell (Reuter et al., 2003) (intra- and inter-assay CVs: 10.1% and 12.2% for arsenobetaine and arsenocholine; 2.7% and 4.7% for As^{III} and As^{V}; 2.8% and 3.9% for MMAs; 0.6% and 1.3% for DMAs). The sum of trivalent and pentavalent forms of each As metabolite are reported due to oxidation during storage. A refractometer was used to measure specific gravity (SG).

Study sample

Eleven participants discontinued the study due to adverse events (N = 6; placebo: abdominal cramps; 400FA: hypertension; 800 FA: abdominal cramps, vertigo, bilateral hydronephrosis; creatine: vertigo), pregnancy (N = 3; 400FA, creatine, and creatine+400FA), and dropout (N = 2; placebo and 400FA). One participant was

dropped due to a missing baseline blood sample.

The current analyses used data from venous blood samples at baseline and week 12 and urinary As metabolites at baseline, week 1, week 6 and week 12. Due to missing baseline samples, two participants were excluded from analyses stratified by choline and betaine status, and one participant was excluded from analyses stratified by vitamin B_{12} and plasma folate status. An additional two participants with missing week 12 blood samples were excluded. A total of 605 participants were available for blood biomarker analyses stratified by baseline choline and betaine status and 606 participants were available for analyses stratified by vitamin B_{12} and plasma folate (placebo: N = 101; 400FA: N = 152; 800FA: N = 149; creatine: N = 100 and 101, respectively; creatine+400FA: N = 103). GAA was measured in a subset of participants to evaluate the effect of creatine supplementation on GAA (Peters et al., 2015b); 400FA and 800FA groups were excluded from analyses of change in GAA.

Missing urine biomarkers or biomarkers associated with missing SG data (N = 5) were excluded from analyses of changes in urinary As metabolites. SG $\leq$ 1.001 is accepted to be outside of the normal range (Vahter et al., 2006). Values $\leq$ 1.001 for SG (N = 48 at baseline; N = 46 at week 12) or %InAs (N = 1 at baseline; N = 5 at week 12) were also excluded (Vahter, 1999). A total of 511 participants were included in analyses of the change in urinary As metabolites stratified by choline and betaine and 512 participants were included in analyses stratified by vitamin B_{12} and plasma folate (placebo: N = 85; 400FA: N = 128; 800FA: N = 122; creatine: N = 87 and 88, respectively; creatine+400FA: N = 89). Samples with urinary As^{III} and As^{V} concentrations below the limit of detection (LOD) were replaced with LOD/2 (0.025 μg/L) (baseline As^{III} N = 5; baseline As^{v} N = 3; week 12 As^{III} N = 6; week 12 As^{V} N = 9). %InAs, %MMAs, and %DMAs in urine were calculated by dividing the concentration of each species by the sum of As^{III} + As^{V} + MMAs + DMAs concentra-

tions. Primary methylation index (PMI) and secondary methylation index (SMI) were calculated (MMAs/InAs and DMAs/MMAs, respectively). Arsenobetaine and arsenocholine were excluded from these calculations because they are non-toxic forms of As from dietary sources (Syracuse Research Corporation 2007).

Statistical analysis

Means and SDs were calculated for baseline characteristics for each treatment group. Differences between treatment groups were assessed using the Chi-square test for categorical variables and the Kruskal-Wallis test for continuous variables. Participants were classified as high or low choline, betaine, plasma folate, and vitamin B_{12} using a median cut-off point (choline: 11.4 nmol/mL; betaine: 43.6 nmol/mL; plasma folate: 13.5 nmol/L; vitamin B_{12}: 214.9 pmol/L). Differences in baseline homocysteine, GAA, bAs, and urinary As metabolite proportions and indices between treatment groups within each nutrient strata were also assessed using the Kruskal-Wallis rank sum test.

The distributions of each outcome (i.e., within-person changes at week 12 in bAs concentration, homocysteine concentration, and urinary As metabolite proportions and indices) and baseline variables were examined. Blood As, homocysteine, %InAs, and SMI were right skewed and natural log-transformation was used to reduce the skewness of the baseline and week 12 variables so that the within-person change met linear model assumptions. Levene's test was used to check the linear model assumption of homoscedasticity.

Analyses of treatment group effects were performed by intent-to-treat. For each nutrient stratum, mean differences between treatment and placebo groups in within-person changes of bAs concentration, homocysteine concentration, or urinary As metabolite proportions and indices were estimated using linear regression models. In the case of heteroscedasticity, standard errors and P-values were calculated by a

heteroscedasticity-consistent covariance matrix estimation using the *sandwich* package in R (Zeileis, 2004). Due to baseline treatment group differences in urinary As metabolite proportions and As methylation indices between strata of choline, betaine, and folate, and in GAA within folate strata (data not shown), models predicting within-person change in urinary As metabolite proportions and As methylation indices stratified by choline, betaine, and folate were adjusted for baseline As metabolite proportions or As methylation indices, respectively. Models predicting change in GAA stratified by folate were adjusted for baseline GAA. A Wald test was used to detect differences between strata (above vs. below median) in the treatment effect. To adjust for multiple tests in detecting differences in treatment effects between strata, the Benjamini-Hochberg adjustment on P-values was used to control for the false discovery rate (FDR) (Benjamini and Hochberg, 1995).

Linear models with repeated measures were used to further examine whether the observed treatment effects on within-person changes in As metabolite proportions over 12 weeks stratified by choline status may be present at weeks 1 or 6 (N = 538 for participants with data on change from baseline to weeks 1, 6 or 12). The natural-log transformation of %InAs at each time point was used. Changes in the proportions of each metabolite since baseline were calculated for weeks 1, 6, and 12. Models included control variable for baseline metabolite proportion, and predictors of treatment group, time categories and group-by-time interactions, which indicate treatment group differences in mean within-person change since baseline. Model parameters were estimated using a generalized estimating equation approach to account for within-subject correlations in the repeated measures.

Analyses were performed using R version 3.2.2 (Vienna, Austria) (R Core Team, 2015) and SAS 9.4 (Cary, NC).

4.4 Results

Baseline participant characteristics are presented in Table 4.1. Participants had a mean age of 38 years (range: 24-55), and approximately half of participants were male (50.5%). There were no significant baseline differences in demographics, As exposure, or nutritional factors between treatment groups ($P > 0.05$). The majority of participants were folate sufficient ($\geq$ 9 nmol/L in plasma: 80.2%) and vitamin B_{12} sufficient ($\geq$ 151 pmol/L: 75.9%).

Treatment effects on homocysteine

There were significant differences in the change in homocysteine over 12 weeks between participants with baseline plasma folate below and above the median (Table 4.2). The mean within-person decrease in ln(homocysteine) relative to placebo was significantly greater in the low folate stratum with 400FA (low folate: B = -0.33, $P <$ 0.001; high folate: B = -0.16, $P < 0.001$; Wald test for difference between strata: $P =$ 0.011, $FDR = 0.024$) and creatine+400FA supplementation (low folate: B = -0.31, P < 0.001; high folate: B = -0.14, $P < 0.002$; Wald test $P = 0.012$, $FDR = 0.024$). The difference between strata in the effects of 800FA on change in ln(homocysteine) was similar but with lower statistical significance (low folate: B = -0.34, $P < 0.001$; high folate: B = -0.20, $P < 0.001$; Wald test $P = 0.049$, $FDR = 0.065$).

Treatment effects on As methylation: FA

When stratifying by baseline betaine, mean within-person decreases in ln(%InAs) relative to placebo were greater among participants below the median with 400FA (low betaine: B = -0.19, $P = 0.009$; high betaine: B = -0.06, $P = 0.28$) and 800FA supplementation (low betaine: B = -0.29, $P < 0.001$; high betaine: B = -0.11, $P = 0.035$) (Table 3). The difference in treatment effects between strata with 800FA was nominally

Table 4.1: Participant characteristics at baseline

	Placebo Mean±SD	**400FA Mean±SD**	**800FA Mean±SD**	**Creatine Mean±SD**	**Creatine+400FA Mean±SD**	***P*[1]**
N	101	152	149	101	103	
Age (years)	37.9±7.3	39.0±8.0	38.1±8.1	38.3±8.2	38.0±7.7	0.85
Male (%)	50.5	50.7	50.3	50.5	50.5	1.00
Smoking ever (%)[2]	24.8	24	29.5	28.7	30.1	0.71
Betel nut use ever (%)[2]	27.7	24	24.2	24.8	20.4	0.81
BMI[4]	20.4±3.1	19.5±2.3	19.8±2.7	20.0±3.0	19.5±2.5	0.24
Water As (μg/L)	158.0±126.4	148.8±117.9	151.1±121.7	158.8±125.4	166.9±146.6	0.94
Blood As (μg/L)	9.8±5.6	11.0±9.8	10.1±5.5	10.7±8.2	10.5±5.3	0.68
Urinary As (μg/L)[5]	171.1±118.4	210.6±259.4	177.7±113.7	183.2±144.7	186.0±106.7	0.5
Urinary %InAs[6]	14.6±4.2	13.9±4.4	13.8±4.5	14.5±5.3	13.2±3.7	0.18
Urinary %MMAs[6]	13.1±4.4	13.2±4.4	13.0±4.9	13.4±4.4	12.5±4.3	0.51
Urinary %DMAs[6]	72.3±7.1	73.0±6.5	73.2±7.5	72.1±7.4	74.4±6.0	0.25
Urinary PMI[6]	0.9±0.3	1.0±0.4	1.0±0.4	1.0±0.4	1.0±0.4	0.59
Urinary SMI[6]	6.3±2.8	6.3±2.5	6.8±4.0	6.2±2.8	6.9±3.0	0.38
Urinary creatinine (mg/dL)[6]	52.0±34.6	61.7±45.2	58.1±41.5	60.0±36.2	65.6±46.8	0.35
Plasma folate (nmol/L)[7]	16.7±17.3	16.6±14.2	17.9±15.9	16.0±7.9	15.4±8.7	0.65
Folate deficient (< 9 nmol/L in plasma) (%)	21.8	23.7	18.1	13.9	20.4	0.38
Plasma homocysteine (μmol/L)	13.9±10.8	13.7±8.8	13.7±10.3	12.4±5.5	12.8±5.6	0.9
Hyperhomocysteinemia (≥ 13μmol/L) (%)	42.6	36.8	39.6	38.6	37.9	0.92
Plasma vitamin B_{12} (pmol/L)	225.7±97.4	246.1±131.1	247.8±142.5	255.9±141.0	236.9±121.0	0.84
Vitamin B_{12} deficient (< 151 pmol/L) (%)	24.8	24.3	26.2	19.8	24.3	0.84
Choline (nmol/mL)[7]	11.5±2.4	11.7±2.7	11.7±2.9	11.9±2.5	11.7±2.5	0.83
Betaine (nmol/mL)[7]	45.9 ± 16.8	45.5 ± 15.1	45.3 ± 18.0	46.8 ± 16.6	45.8 ± 18.2	0.94
Plasma GAA	2.05±0.66	-	-	1.95±0.57	1.98±0.67	0.16

1. Kruskal-Wallis test and Chi-squared test for group differences in continuous variables and categorical variables, respectively. 2. 400FA: N = 150. 2. 400FA: N = 150. 3. Creatine+400FA: N = 102. 4. Placebo: N = 100, 400FA: N = 149; 800FA: N = 146; Creatine: N = 98; Creatine+400FA: N = 102. 5. Adjusted for SG to account for between-individual differences in urine dilution. Placebo: N = 94, 400FA: N = 140; 800FA: N = 133; Creatine: N = 93; Creatine+400FA: N = 97. 6. Placebo: N = 94, 400FA: N = 140; 800FA: N = 133; Creatine: N = 94; Creatine+400FA: N = 97. 7. Creatine: N = 100. 9. GAA was measured in a subset of participants; 400FA and 800FA groups were from analyses. Creatine+400FA: N = 102. Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg treatment group; PMI, primary methylation index; SMI, secondary methylation index.

Table 4.2: Linear models for change in ln(homocysteine) over 12 weeks, by baseline plasma folate strata.[1]

Change[2]	Low strata (< median)		High strata (> median)		Wald test[3]	
	***B* (95% CI)**	***P***	***B* (95% CI)**	***P***	***P***	***FDR***
400FA	-0.33 (-0.43, -0.23)	<0.001	-0.16 (-0.25, -0.08)	<0.001	0.011	0.024
800FA	-0.34 (-0.45, -0.23)	<0.001	-0.20 (-0.29, -0.11)	<0.001	0.049	0.065
Creatine	-0.05 (-0.15, 0.06)	0.38	-0.02 (-0.11, 0.07)	0.67	0.7	0.7
Creatine+400FA	-0.31 (-0.42, -0.21)	<0.001	-0.14 (-0.22, -0.05)	0.002	0.012	0.024

1. Placebo used as reference group. Baseline plasma folate median = 13.50 nmol/L. 2. Week 12 - week 1, treatment vs. placebo. 3. Test for difference between strata.
Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group.

significant (Wald test $P = 0.04$, $FDR = 0.17$). The decrease in %MMAs relative to placebo was greater among participants in the low betaine stratum with 400FA (low betaine: B = -3.06, $P < 0.001$; high betaine: B = -1.35, $P = 0.01$; Wald test $P = 0.03$, $FDR = 0.10$). Significantly greater mean within-person increases in %DMAs relative to placebo were also observed in the low betaine strata with 400FA (low betaine: B = 1.02, $P < 0.001$; high betaine: $B = 1.37$, $P = 0.07$; Wald test $P = 0.011$, FDR = 0.044) and 800FA (low betaine: $B = 7.07$, $P < 0.001$; high betaine: $B = 3.49$, P = 0.001; Wald test $P = 0.022$, $FDR = 0.044$). Differences in 400FA treatment effects between betaine strata were reflected in a greater mean within-person increase in ln(SMI) relative to placebo in the low stratum ($B = 0.39$, $P < 0.001$) than the high stratum (B = 0.14, $P = 0.001$) (Wald test $P = 0.005$; $FDR = 0.021$). Linear models with repeated measures indicated an increasing effect size of 400FA and 800FA on the change in As metabolite proportions at weeks 1, 6, and 12 (Appendix Table 4.5).

Treatment effects on As methylation: Creatine

We observed differences in creatine treatment effects on the mean within-person changes in As metabolites proportions over 12 weeks between participants above

Table 4.3: Linear models for change in As metabolite proportions over 12 weeks, by baseline choline strata.[1]

Change[2]	Low strata (≤ median)		High strata (> median)		Wald test[3]	
	***B* (95% CI)**	***P***	***B* (95% CI)**	***P***	***P***	***FDR***
ln(%InAs)[4]						
400FA	-0.12 (-0.24, 0.01)	0.078	-0.12 (-0.24, 0.00)	0.048	0.955	0.955
800FA	-0.15 (-0.27, -0.03)	0.017	-0.23 (-0.34, -0.12)	<0.001	0.344	0.688
Creatine	0.05 (-0.09, 0.19)	0.466	-0.11 (-0.22, 0.00)	0.053	0.077	0.306
Creatine+400FA	-0.12 (-0.30, 0.06)	0.185	-0.14 (-0.25, -0.03)	0.012	0.855	0.955
%MMAs[5]						
400FA	-3.00 (-4.09, -1.91)	<0.001	-1.25 (-2.25, -0.25)	0.014	0.02	0.078
800FA	-3.23 (-4.34, -2.12)	<0.001	-1.76 (-2.72, -0.80)	<0.001	0.049	0.097
Creatine	-1.47 (-2.88, -0.07)	0.04	0.05 (-1.09, 1.18)	0.936	0.098	0.108
Creatine+400FA	-2.91 (-4.11, -1.72)	<0.001	-1.55 (-2.72, -0.38)	0.01	0.108	0.108
%DMAs[6]						
400FA	4.53 (2.03, 7.04)	<0.001	2.71 (0.71, 4.72)	0.008	0.265	0.931
800FA	5.54 (3.21, 7.88)	<0.001	4.58 (2.72, 6.44)	<0.001	0.524	0.931
Creatine	0.93 (-1.98, 3.83)	0.53	1.53 (-0.41, 3.47)	0.121	0.734	0.931
Creatine+400FA	3.10 (-0.24, 6.44)	0.069	3.27 (1.35, 5.19)	0.001	0.931	0.931

1. Placebo used as reference group. Baseline choline median = 11.42 nmol/mL. 2. Week 12 - week 1, treatment vs. placebo. 3. Test for difference between strata. 4. Adjusted for baseline ln(%InAs). 5. Adjusted for baseline %MMAs. 6. Adjusted for baseline %DMAs. Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group; InAs, inorganic arsenic; MMAs, monomethyl-arsenical species; DMAs; dimethyl-arsenical species.

and below the median baseline choline (Table 4.4 and Figure 4.3). Creatine treatment led to a significant decrease in urinary %MMAs compared to placebo among participants in the low choline stratum (B = -1.47, P = 0.04), but not among participants in the high stratum (B = 0.05, P = 0.94), although the difference in the treatment effect was not significant (Wald test P = 0.10). While there were no clear creatine effects on %InAs or %DMA, the three percentages are interrelated and changes in PMI and SMI were analyzed to evaluate overall direction of creatine treatment effects on As methylation. There was a significant mean within-person decrease in PMI and increase in ln(SMI) among participants in the low choline stratum (PMI: B = -0.16, P = 0.035; ln(SMI): B = 0.14, P = 0.035), but not among participants in the high stratum (PMI:

B = 0.06, P = 0.37; ln(SMI): B = 0.02, P = 0.69). Effect sizes for change in PMI differed between strata, but the statistical significance was marginal after correcting for multiple tests (Wald test P = 0.028; FDR = 0.06). A similar pattern was observed when stratifying by betaine: the increase in ln(SMI) was significant in the low betaine stratum (B = 0.14, P = 0.033), but not the high betaine stratum (B = 0.04, P = 0.46; Wald test P = 0.26) (Appendix Table 4.6).

In linear models with repeated measures, the mean within-person decrease in %MMAs was significantly greater in the creatine group than the placebo group among participants in the low choline stratum beginning after one week of supplementation (P = 0.028), and remained significantly greater at weeks 6 (P = 0.003) and 12 (P = 0.022) (Appendix Table 4.5). However, mean within-person changes in %MMAs were not significantly different between the creatine and placebo groups in the high choline stratum at any follow-up point.

Treatment effects on GAA: Creatine.

The mean within-person decrease in ln(GAA) with creatine treatment relative to placebo was significant in the high choline and plasma folate strata (high choline: B = -0.21, $P < 0.001$; high folate: B = -0.22, $P < 0.001$) (Appendix Table 4.6). The treatment effect was not significant in the low strata for either nutrient, although it was suggestive in the low folate stratum (low folate stratum: B = -0.10, P = 0.052). The difference in treatment effects by strata were marginally significant before correction for multiple tests (Wald test P = 0.083 for difference between choline strata; P = 0.091 for difference between folate strata).

Complete results for regression analyses stratified by baseline choline, betaine, vitamin B_{12}, and plasma folate concentrations for all treatment groups are presented in Appendix Table 1. We did not observe differences in the mean within-person change

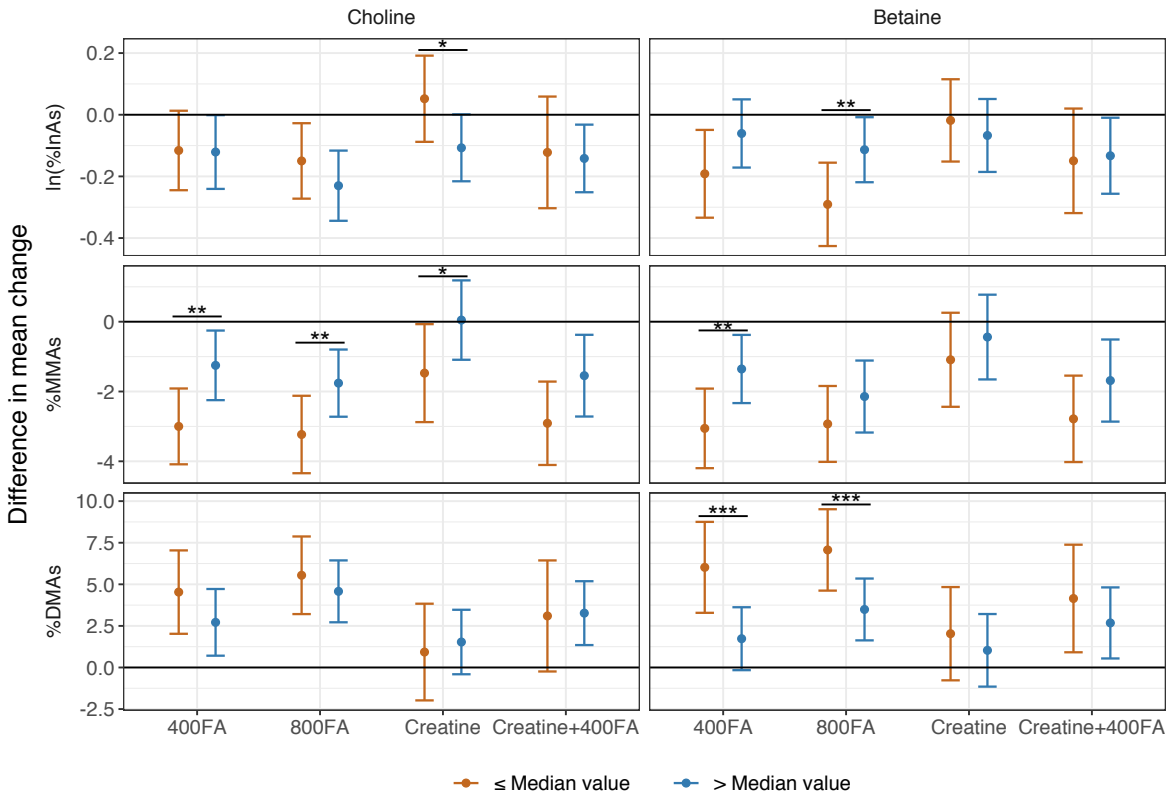


Figure 4.3: Differences in mean change in urinary As metabolite proportions (week 12 − week 0) between treatment and placebo stratified by baseline choline and betaine below and above median. *P*-values are from Wald test for differences between strata in treatment effects based on linear models for change in As metabolite proportions adjusting for baseline levels of As metabolite proportions, Baseline choline median = 11.42 nmol/mL; baseline betaine median = 43.63 nmol/mL. * Wald test for differences between strata $P < 0.10$; ** $P < 0.05$; *** $FDR < 0.05$. Abbreviations: 400FA, 400 µg FA/day treatment group; 800FA, 800 µg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 µg FA/day treatment group; InAs, inorganic arsenic; MMAs, monomethyl-arsenical species; DMAs; dimethyl-arsenical species.

Table 4.4: Linear models for change in As metabolite proportions over 12 weeks, by baseline betaine strata.[1]

Change[2]	Low strata (≤ median)		High strata (>median)		Wald test[3]	
	***B* (95% CI)**	***P***	***B* (95% CI)**	***P***	***P***	***FDR***
ln(%InAs)[4]						
400FA	-0.19 (-0.33, -0.05)	0.009	-0.06 (-0.17, 0.05)	0.281	0.153	0.306
800FA	-0.29 (-0.43, -0.16)	<0.001	-0.11 (-0.22, -0.01)	0.035	0.042	0.167
Creatine	-0.02 (-0.15, 0.12)	0.787	-0.07 (-0.19, 0.05)	0.264	0.589	0.786
Creatine+400FA	-0.15 (-0.32, 0.02)	0.084	-0.13 (-0.26, -0.01)	0.035	0.877	0.877
%MMAs[5]						
400FA	-3.06 (-4.20, -1.92)	<0.001	-1.35 (-2.33, -0.38)	0.007	0.026	0.102
800FA	-2.93 (-4.02, -1.84)	<0.001	-2.14 (-3.18, -1.11)	<0.001	0.302	0.403
Creatine	-1.09 (-2.44, 0.26)	0.112	-0.44 (-1.65, 0.77)	0.476	0.48	0.48
Creatine+400FA	-2.78 (-4.02, -1.55)	<0.001	-1.69 (-2.86, -0.51)	0.005	0.206	0.403
%DMAs[6]						
400FA	6.02 (3.29, 8.75)	<0.001	1.73 (-0.16, 3.62)	0.073	0.011	0.044
800FA	7.07 (4.62, 9.51)	<0.001	3.49 (1.63, 5.35)	<0.001	0.022	0.044
Creatine	2.03 (-0.77, 4.83)	0.155	1.03 (-1.15, 3.21)	0.353	0.579	0.579
Creatine+400FA	4.15 (0.91, 7.38)	0.012	2.68 (0.54, 4.81)	0.014	0.455	0.579

1. Placebo used as reference group. Baseline betaine median = 43.63 nmol/mL. 2. Week 12 - week 1, treatment vs. placebo. 3. Test for difference between strata. 4. Adjusted for baseline ln(%InAs). 5. Adjusted for baseline %MMAs. 6. Adjusted for baseline %DMAs. Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group; InAs, inorganic arsenic; MMAs, monomethyl-arsenical species; DMAs; dimethyl-arsenical species.

in ln(bAs) with FA or creatine supplementation between the high and low strata of choline, betaine, vitamin B_{12}, or plasma folate. In addition, the mean within-person changes in ln(homocysteine), ln(bAs), or urinary As metabolite proportions with FA or creatine treatment relative to placebo did not differ by vitamin B_{12} strata.

4.5 Discussion

Arsenic methylation capacity is influenced by OCM. Folate plays an important role in recruiting one-carbon units for the remethylation of homocysteine and the synthesis of SAM. The availability of one-carbon units is also impacted by other nu-

trients including the alternative methyl donor betaine, its precursor choline (Chiuve et al., 2007; Holm et al., 2005), and possibly the cofactor vitamin B_{12}. This study investigated whether these nutrients modify FA and creatine treatment effects on changes in total homocysteine and GAA concentrations (biomarkers of OCM and endogenous creatine synthesis, respectively), total bAs concentrations, and urinary As metabolite proportions and methylation indices.

Treatment effects on homocysteine

In agreement with previous studies of the homocysteine-lowering effects of FA treatment (Homocysteine Lowering Trialists' Collaboration, 1998), we observed a greater mean decrease in homocysteine concentration among participants with lower baseline plasma folate concentrations.

Treatment effects on As metabolites

We observed that the effect of FA supplementation on the change in As methylation capacity was more pronounced among participants with baseline plasma betaine concentrations below the median. 400FA was associated with significant changes in urinary %InAs and %DMAs among participants in the low betaine stratum, but not among participants in the high betaine stratum. The mean within-person increase in %DMAs with 800FA relative to placebo was also significantly greater among participants in the low betaine stratum compared with the high betaine stratum. These observations support our hypothesis that FA treatment effects would be greater among those with nutritional deficiencies. The complementary role of folate for the remethylation of homocysteine under conditions of low betaine is well characterized (Obeid, 2013), and the effects of this on increasing As methylation are relatively straightforward to interpret.

Urinary creatinine has been associated with lower %InAs and higher %DMAs in previous cross-sectional analyses (Basu et al., 2011; Bozack et al., 2018a; Gamble et al., 2005, 2006; Hall et al., 2009, 2007; Kile et al., 2009; Pilsner et al., 2009), consistent with results from cross-sectional analyses at baseline in the current study. Creatine supplementation was associated with a mean within-person decrease in %MMAs overall at weeks 1, 6 and 12, (as previously reported) (Bozack et al., 2018a). Here we find that compared to placebo group, creatine supplementation was associated with a rapid decrease in %MMAs at week 1 that then plateaus at weeks 6 and 12 among participants with choline concentrations below the median. Creatine treatment was not associated with a change in %MMAs among participants with choline above the median. Among numerous other roles, choline serves as a precursor in the biosynthesis of betaine, an alternative methyl donor in OCM. Plasma choline and betaine concentrations are significantly correlated (Holm et al., 2003; Nurk et al., 2012), including in the current study ($r_{Spearman} = 0.45$; $P < 0.001$ at baseline). Both synthesis of PC and creatine are major consumers of SAM (Figure 2.4). Low choline status stimulates PC synthesis from phosphatidylethanolamine (PE) via PE methyl transferase (PEMT), requiring 3 molecules of SAM per molecule of PC (Kennedy, 1957). While the resulting PC can be converted to choline and then betaine, much PC is needed to satisfy other essential roles of choline and PC in processes such as lipid transport, cell signaling and maintaining the structural integrity of cell membranes. We speculate that when choline concentrations are low and PEMT is upregulated, more SAM may be used in the PC synthesis, resulting in a low-SAM scenario, e.g., below the KM of AS3MT for SAM, allowing a more pronounced impact of creatine supplementation in As methylation. In support of this hypothesis, SAM concentrations in whole blood were lower in the low choline strata than in high choline strata (t-test $P = 0.017$), though we do not know how well this reflects hepatic SAM concentrations. Possibly, the cross-sectional relationships

between urinary creatinine and %InAs and %DMAs are related to OCM but also in part to renal tubular reabsorption of InAs under conditions of more concentrated urine (Ginsburg and Lotspeich, 1963), whereas creatine treatment effects are directly related to liver-specific OCM alterations that are attenuated over time by long-range allosteric regulation of hepatic SAM concentrations. The relationships between choline status and creatine supplementation are likely complex, and achieving a fuller understanding of these findings requires further study.

We did not observe effect modification of FA or creatine treatment on the change in As methylation by vitamin B_{12} status. Vitamin B_{12} serves as a cofactor for the remethylation of homocysteine, and vitamin B_{12} deficiency limits the availability of one-carbon units carried by 5-methyl-THF for the synthesis of SAM (Green et al., 2017). Although one might expect that vitamin B_{12} status would modify treatment effects on As methylation capacity, evidence of the association between vitamin B_{12} and As methylation capacity is inconsistent (Hall et al., 2009; López-Carrillo et al., 2016; Spratlen et al., 2017). In addition, this may not have been detected due to the relatively low prevalence of vitamin B_{12} deficiency among study participants.

Similarly, we did not observe effect modification of FA or creatine treatment on the change in As methylation by baseline plasma folate status. This suggests that FA supplementation enhances As methylation capacity even among individuals with high folate status. This finding may also be due to elevated homocysteine in this population. Although 38.9% of study participants had hyperhomocysteinemia ($\geq$ 13 μmol/L) at baseline, 63% of women and 72.9% of men had homocysteine levels above the normal range defined by the U.S. CDC (4.5-7.9 μmol/L for women and 6.3-11.2 μmol/L for men) (Centers for Disease Control and Prevention, 2003). Given this high prevalence of elevated homocysteine, FA treatment may have the effect of lowering homocysteine and SAH even among participants classified as having plasma folate above the median.

Treatment effects on guanidinoacetate

Creatine supplementation pre-translationally inhibits AGAT to downregulate the first step of creatine synthesis (Figure 2). As previously reported, creatine supplementation lowered GAA concentrations (a product of AGAT) in the overall study population; as expected, FA did not affect GAA (Peters et al., 2015b). However, significant treatment effects of creatine in lowering GAA were observed in the high strata of choline and plasma folate but not in the low strata. This may reflect reduced statistical power in stratified sub-samples. However, an alternative explanation is that the hepatic methylation of GAA that exists prior to creatine supplementation is not inhibited by creatine supplementation (da Silva et al., 2009) and may be more active in a folate replete state when hepatic SAM concentrations are expected to be relatively high. It should be noted that sample size limited our power to identify effects when stratifying by baseline nutritional status. The sample size for this RCT was originally selected to identify our primary outcomes of total treatment group effects. In addition, we selected the median baseline value to categorize participants as having high or low baseline nutritional status. This approach was reasonable for choline and betaine, micronutrients for which there are no reference ranges to determine deficiency, to maximize power. However, we were not able to determine if threshold effects could result in effect modification at different baseline micronutrient levels.

4.6 Conclusion

This study contributes to an understanding of the relationship between FA and creatine supplementation, OCM-related micronutrients, and As methylation capacity, and their dependence on nutritional status. Although removal of As from drinking water is the primary and most effective approach to decreasing As-related morbidity

and mortality (World Health Organization, 2012), As exposure remains a persistent public health concern in many regions of the world (Naujokas et al., 2013), including the U.S. (Mantha et al., 2017). Our group has previously reported the effects of FA and creatine supplementation on total bAs concentrations and As methylation capacity (Bozack et al., 2018a; Gamble et al., 2006, 2007; Peters et al., 2015a). We have observed that FA supplementation significantly lowers bAs and increases As methylation capacity, and creatine supplementation decreases the urinary proportion of MMAs. Interestingly, FA treatment effects on As methylation did not differ by baseline folate status, though the effects of FA and creatine supplementation on As methylation capacity were greater among individuals with low choline and betaine status. These observations are particularly relevant among As-exposed individuals who may or may not be folate deficient but who have sub-optimal choline status, including populations in the U.S. (Zeisel, 2009). Although further research is necessary to fully understand the relationship between OCM-related micronutrients and As metabolism, these findings highlight the potential of nutritional interventions to reduce the adverse health effects of As exposure, particularly in regions with a high prevalence of nutrient deficiencies.

4.7 Bibliography

Ahsan, H; Chen, Y; Parvez, F; Argos, M; Hussain, A I; Momotaj, H; Levy, D; van Geen, A; Howe, G, and Graziano, J. Health Effects of Arsenic Longitudinal Study (HEALS): Description of a multidisciplinary epidemiologic investigation. *Journal of Exposure Science and Environmental Epidemiology*, 16(2):191–205, 2006. doi: 10.1038/sj.jea.7500449.

Bangladesh Bureau of Statistics and United Nation Children's Fund. Bangladesh multiple indicator cluster survey 2012-2013 final report. Technical report, Dhaka, Bangladesh, 2015.

Basu, A; Mitra, S; Chung, J; Guha Mazumder, DN; Ghosh, N; Kalman, D; von Ehrenstein, O S; Steinmaus, C; Liaw, J, and Smith, A H. Creatinine, diet, micronutrients, and arsenic methylation in West Bengal, India. *Environmental Health Perspectives*, 119(9):1308–1313, 2011. doi: 10.1289/ehp.1003393.

Benbrahim-Tallaa, L and Waalkes, M P. Inorganic arsenic and human prostate cancer. *Environmental Health Perspectives*, 116(2):158–64, 2008. doi: 10.1289/ehp.10423.

Benjamini, Y and Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57 (1):289–300, 1995.

Bozack, A K; Hall, M N; Liu, X; Ilievski, V; Lomax-Luu, A M; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T, and Graziano, J H. Folic acid supplementation enhances arsenic methylation: results from a folic acid and creatine supplementation randomized controlled trial in Bangladesh. *American Journal of Clinical Nutrition*, 109(2):380–391, 2018a. doi: 10.1093/ajcn/nqy148.

Bozack, A K; Saxena, R, and Gamble, M V. Nutritional influences on one-carbon metabolism: effects on arsenic methylation and toxicity. *Annual Review of Nutrition*, 38:401–429, 2018b. doi: 10.1146/annurev-nutr-082117-051757.

Brosnan, J T; da Silva, R P, and Brosnan, M E. The metabolic burden of creatine synthesis. *Amino Acids*, 40(5):1325–31, 2011. doi: 10.1007/s00726-011-0853-y.

Carducci, C; Birarelli, M; Leuzzi, V; Carducci, C; Battini, R; Cioni, G, and Antonozzi, I. Guanidinoacetate and creatine plus creatinine assessment in physiologic fluids: an effective diagnostic tool for the biochemical diagnosis of arginine:glycine amidinotransferase and guanidinoacetate methyltransferase deficiencies. *Clinical Chemistry*, 48(10):1772–8, 2002.

Centers for Disease Control and Prevention. Laboratory Procedure Manual: Total Homocysteine (tHcy), NHANES 2003-2004. Technical report, Atlanta, GA, 2003. URL `https://www.cdc.gov/nchs/data/nhanes/nhanes_03_04/l06_c_met_homocysteine.pdf`.

Challenger, F. Biological methylation. *Chemical Reviews*, 36(3):315–361, 1945. doi: 10.1021/cr60115a003.

Chiuve, S E; Giovannucci, E L; Hankinson, S E; Zeisel, S H; Dougherty, L W; Willett, W C, and Rimm, E B. The association between betaine and choline intakes and the plasma concentrations of homocysteine in women. *American Journal of Clinical Nutrition*, 86(4):1073–1081, 2007. doi: 10.1093/ajcn/86.4.1073.

da Silva, R P; Nissim, I; Brosnan, M E, and Brosnan, J T. Creatine synthesis: hepatic metabolism of guanidinoacetate and creatine in the rat *in vitro* and *in vivo*. *American Journal of Physiology*, 296(2):E256–61, 2009. doi: 10.1152/ajpendo.90547.2008.

Gamble, M V; Liu, X; Ahsan, H; Pilsner, R; Ilievski, V; Slavkovich, V; Parvez, F; Levy, D; Factor-Litvak, P, and Graziano, J H. Folate, homocysteine, and arsenic metabolism in arsenic-exposed individuals in Bangladesh. *Environmental Health Perspectives*, 113(12):1683–8, 2005. doi: 10.1289/ehp.8084.

Gamble, M V; Liu, X; Ahsan, H; Pilsner, J R; Ilievski, V; Slavkovich, V; Parvez, F; Chen, Y; Levy, D; Factor-Litvak, P, and Graziano, J H. Folate and arsenic metabolism: a double-blind, placebo-controlled folic acidsupplementation trial in Bangladesh. *American Journal of Clinical Nutrition*, 84(5):1093–101, 2006. doi: 10.1093/ajcn/84.5.1093.

Gamble, M V; Liu, X; Slavkovich, V; Pilsner, J R; Ilievski, V; Factor-Litvak, P; Levy, D; Alam, S; Islam, M; Parvez, F; Ahsan, H, and Graziano, J H. Folic acid supplementation lowers blood arsenic. *American Journal of Clinical Nutrition*, 86(4):1202–9, 2007. doi: 10.1093/ajcn/86.4.1202.

Ginsburg, J M and Lotspeich, W D. Interrelations of arsenate and phosphate transport in the dog kidney. *American Journal of Physiology*, 205(4):707–714, 1963. doi: 10.1152/ajplegacy.1963.205.4.707.

Green, R; Allen, L H; Bjørke-Monsen, A-L; Brito, A; Guéant, J-L; Miller, J W; Molloy, A M; Nexo, E; Stabler, S; Toh, B-H; Ueland, P M, and Yajnik, C. Vitamin B_{12} deficiency. *Nature Reviews Disease Primers*, 3:17040, 2017. doi: 10.1038/nrdp.2017.40.

Hall, M; Chen, Y; Ahsan, H; Slavkovich, V; van Geen, A; Parvez, F, and Graziano, J H. Blood arsenic as a biomarker of arsenic exposure: results from a prospective study. *Toxicology*, 225(2-3):225–33, 2006. doi: 10.1016/j.tox.2006.06.010.

Hall, M; Gamble, M; Slavkovich, V; Liu, X; Levy, D; Cheng, Z; van Geen, A; Yunus, M; Rahman, M; Pilsner, J R, and Graziano, J. Determinants of arsenic metabolism: blood arsenic metabolites, plasma folate, cobalamin, and homocysteine concentrations in maternalnewborn pairs. *Environmental Health Perspectives*, 115(10):1503–1509, 2007. doi: 10.1289/ehp.9906.

Hall, M N; Liu, X; Slavkovich, V; Ilievski, V; Pilsner, J R; Alam, S; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate, cobalamin, cysteine, homocysteine, and arsenic metabolism among children in Bangladesh. *Environmental Health Perspectives*, 117(5), 2009. doi: 10.1289/ehp.0800164.

Hall, M N; Howe, C G; Liu, X; Caudill, M A; Malysheva, O; Ilievski, V; Lomax-Luu, A M; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T; Graziano, J H, and Gamble, M V. Supplementation with folic acid, but not creatine, increases plasma betaine, decreases plasma dimethylglycine, and prevents a decrease in plasma choline in arsenic-exposed Bangladeshi adults. *Journal of Nutrition*, 146 (5):1062–1067, 2016. doi: 10.3945/jn.115.227132.

Heck, J E; Gamble, M V; Chen, Y; Graziano, J H; Slavkovich, V; Parvez, F; Baron, J A; Howe, G R, and Ahsan, H. Consumption of folate-related nutrients and metabolism of arsenic in Bangladesh. *American Journal of Clinical Nutrition*, 85(5):1367–74, 2007. doi: 10.1093/ajcn/85.5.1367.

Holm, P I; Ueland, P M; Kvalheim, G, and Lien, E A. Determination of choline, betaine, and dimethylglycine in plasma by a high-throughput method based on normal-phase chromatography-tandem mass spectrometry. *Clinical Chemistry*, 49(2):286–94, 2003. doi: 10.1373/49.2.286.

Holm, P I; Ueland, P M; Vollset, S Emil; Midttun, Ø; Blom, H J; Keijzer, Miranda B A J, and den Heijer, M. Betaine and folate status as cooperative determinants of plasma homocysteine in humans. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 25(2):379–85, 2005. doi: 10.1161/01.ATV.0000151283.33976.e6.

Homocysteine Lowering Trialists' Collaboration. Lowering blood homocysteine with folic acid based supplements: meta-analysis of randomised trials. Homocysteine Lowering Trialists' Collaboration. *BMJ*, 316(7135):894–8, 1998. doi: 10.1136/BMJ.316.7135.894.

Howe, C G; Liu, X; Hall, M N; Ilievski, V; Caudill, M A; Malysheva, O; Lomax-Luu, A M; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T; Graziano, J H; Costa, M, and Gamble, M V. Sex-specific associations between one-carbon metabolism indices and posttranslational histone modifications in arsenic-exposed Bangladeshi adults. *Cancer Epidemiology, Biomarkers and Prevention*, 26 (2):261–269, 2017. doi: 10.1158/1055-9965.EPI-16-0202.

Kennedy, E P. Metabolism of lipides. *Annual Review of Biochemistry*, 26(1):119–148, 1957. doi: 10.1146/annurev.bi.26.070157.001003.

Kile, M L; Hoffman, E; Hsueh, Y-M; Afroz, S; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Ryan, L, and Christiani, D C. Variability in biomarkers of arsenic exposure and metabolism in adults over time. *Environmental Health Perspectives*, 117(3):455–460, 2009. doi: 10.1289/ehp.11251.

Kuo, C-C; Moon, K A; Wang, S-L; Silbergeld, E, and Navas-Acien, A. The association of arsenic metabolism with cancer, cardiovascular disease, and diabetes: a systematic review of the epidemiological evidence. *Environmental Health Perspectives*, 125(8), 2017. doi: 10.1289/EHP577.

López-Carrillo, L; Gamboa-Loira, B; Becerra, W; Hernández-Alcaraz, C; Hernández-Ramírez, R U; Gandolfi, A J; Franco-Marina, F, and Cebrián, M E. Dietary micronutrient intake and its relationship with arsenic metabolism in Mexican women. *Environmental Research*, 151:445–450, 2016. doi: 10.1016/j.envres.2016.08.015.

Mantha, M; Yeary, E; Trent, J; Creed, P A; Kubachka, K; Hanley, T; Shockey, N; Heitkemper, D; Caruso, J; Xue, J; Rice, G; Wymer, L, and Creed, J T. Estimating inorganic arsenic exposure from U.S. rice and total water intakes. *Environmental Health Perspectives*, 125(5):057005, 2017. doi: 10.1289/EHP418.

McGuire, D M; Gross, M D; Van Pilsum, J F, and Towle, H C. Repression of rat kidney L-arginine:glycine amidinotransferase synthesis by creatine at a pretranslational level. *The Journal of Biological Chemistry*, 259(19):12034–8, 1984.

Moe, B; Peng, H; Lu, X; Chen, B; Chen, L WL; Gabos, S; Li, X-F, and Le, X C. Comparative cytotoxicity of fourteen trivalent and pentavalent arsenic species determined using real-time cell sensing. *Journal of Environmental Sciences*, 49:113–124, 2016. doi: 10.1016/j.jes.2016.10.004.

Naujokas, M F.; Anderson, B; Ahsan, H; Vasken Aposhian, H; Graziano, J H.; Thompson, C, and Suk, W A. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environmental Health Perspectives*, 121(3):295–302, 2013. doi: 10.1289/ehp.1205875.

Niculescu, M D and Zeisel, S H. Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. *Journal of Nutrition*, 132(8):2333S–2335S, 2002. doi: 10.1093/jn/132.8.2333S.

Nurk, E; Refsum, H; Bjelland, I; Drevon, C A; Tell, G S; Ueland, P M; Vollset, S E; Engedal, K; Nygaard, H A, and Smith, A D. Plasma free choline, betaine and cognitive performance: the Hordaland Health Study. *British Journal of Nutrition*, 109(3)::511–9, 2012.

Obeid, R. The metabolic burden of methyl donor deficiency with focus on the betaine homocysteine methyltransferase pathway. *Nutrients*, 5(9):3481–95, 2013. doi: 10.3390/nu5093481.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Sanchez, T; van Geen, A; Mey, J L; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T; Balac, O; Ilievski, V; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folic acid and creatine as therapeutic approaches to lower blood arsenic: a randomized controlled trial. *Environmental Health Perspectives*, 123(12):1294–1301, 2015a. doi: 10.1289/ehp.1409396.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M Nasir; Islam, T; Ilievski, V; Graziano, J H, and Gamble, M V. Low-dose creatine supplementation lowers plasma guanidinoacetate, but not plasma homocysteine, in a double-blind, randomized, placebo-controlled trial. *Journal of Nutrition*, 145(10): 2245–52, 2015b. doi: 10.3945/jn.115.216739.

Petrick, J S; Ayala-Fierro, F; Cullen, W R; Carter, D E, and Vasken Aposhian, H. Monomethylarsonous acid (MMA(III)) is more toxic than arsenite in Chang human hepatocytes. *Toxicology and Applied Pharmacology*, 163(2):203–207, 2000. doi: 10.1006/taap.1999.8872.

Pfeiffer, C M; Huff, D L, and Gunter, E W. Rapid and accurate HPLC assay for plasma total homocysteine and cysteine in a clinical laboratory setting. *Clinical chemistry*, 45(2):290–2, 1999.

Pilsner, J R; Liu, X; Ahsan, H; Ilievski, V; Slavkovich, V; Levy, D; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. *Environmental Health Perspectives*, 117(2):254–60, 2009. doi: 10.1289/ehp.11872.

R Core Team. R: A Language and Environment for Statistical Computing. Technical report, R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `https://www.r-project.org/`.

Reuter, W; Davidowski, L, and Neubauer, K. Speciation of five arsenic compounds in urine by HPLC/ICP-MS, PerkinElmerSCIEX. Technical report, PerkinElmer Life and Analytical Sciences, 2003.

Sanchez, T R; Levy, D; Shahriar, M H; Uddin, M N; Siddique, A B; Graziano, J H; Lomax-Luu, A; van Geen, A, and Gamble, M V. Provision of well-water treatment units to 600 households in Bangladesh: a longitudinal analysis of urinary arsenic indicates fading utility. *Science of The Total Environment*, 563:131–137, 2016. doi: 10.1016/j.scitotenv.2016.04.112.

Spratlen, M J; Gamble, M V; Grau-Perez, M; Kuo, C-C; Best, L G; Yracheta, J; Francesconi, K; Goessler, W; Mossavar-Rahmani, Y; Hall, M; Umans, J G; Fretts, A, and Navas-Acien, A. Arsenic metabolism and one-carbon metabolism at low-moderate arsenic exposure: evidence from the Strong Heart Study. *Food and Chemical Toxicology*, 105:387–397, 2017. doi: 10.1016/j.fct.2017.05.004.

Steinmaus, C; Yuan, Y; Kalman, D; Rey, O A; Skibola, C F; Dauphine, D; Basu, A; Porter, K E; Hubbard, A; Bates, M N; Smith, M T, and Smith, A H. Individual differences in arsenic metabolism and lung cancer in a case-control study in Cordoba, Argentina. *Toxicology and Applied Pharmacology*, 247(2):138–45, 2010. doi: 10.1016/j.taap.2010.06.006.

Thomas, D; Waters, Stephen B, and Styblo, Miroslav. Elucidating the pathway for arsenic methylation. *Toxicology and Applied Pharmacology*, 198(3):319–326, 2004. doi: 10.1016/j.taap.2003.10.020.

Tice, R R; Yager, J W; Andrews, P, and Crecelius, E. Effect of hepatic methyl donor status on urinary excretion and DNA damage in B6C3F1 mice treated with sodium arsenite. *Mutation Research*, 386(3):315–34, 1997. doi: 10.1016/S1383-5742(97)00004-5.

Vahter, M. Methylation of inorganic arsenic in different mammalian species and population groups. *Science progress*, 82:69–88, 1999.

Vahter, M and Marafante, E. Effects of low dietary intake of methionine, choline or proteins on the biotransformation of arsenite in the rabbit. *Toxicology Letters*, 37(1):41–46, 1987. doi: 10.1016/0378-4274(87)90165-2.

Vahter, M E; Li, L; Nermell, B; Rahman, A; El Arifeen, S; Rahman, M; Persson, L A, and Ekström, E-C. Arsenic exposure in pregnancy: a population-based study in Matlab, Bangladesh. *Journal of Health, Population, and Nutrition*, 24(2):236–45, 2006.

World Health Organization. Arsenic, 2012. URL `http://www.who.int/mediacentre/factsheets/fs372/en/`.

Yan, J; Jiang, X; West, A A; Perry, C A; Malysheva, O V; Devapatla, S; Pressman, E; Vermeylen, F; Stabler, S P; Allen, R H, and Caudill, M A. Maternal choline intake modulates maternal and fetal biomarkers of choline metabolism in humans. *American Journal of Clinical Nutrition*, 95(5):1060–1071, 2012. doi: 10.3945/ajcn.111.022772.

Zeileis, Achim. Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, 11(10):1–17, 2004.

Zeisel, S H. Importance of methyl donors during reproduction. *American Journal of Clinical Nutrition*, 89(2), 2009. doi: 10.3945/ajcn.2008.26811D.

4.8 Appendix

Table 4.5: Linear models with repeated measures for mean within-person change in As metabolite proportions at weeks 1, 6, and 12 stratified by median baseline choline.[1]

			Low strata (≤ median)		**High strata (> median)**	
Arsenic metabolite	**Treatment group comparison**	**Week**	**Mean difference**	***P***	**Mean difference**	***P***
ln(%InAs)	FA400 vs. placebo	Week 1	-0.04	0.55	-0.07	0.36
		Week 6	-0.08	0.21	-0.21	0.004
		Week 12	-0.12	0.051	-0.12	0.043
	FA800 vs. placebo	Week 1	-0.1	0.19	-0.16	0.021
		Week 6	-0.05	0.44	-0.27	<0.001
		Week 12	-0.15	0.014	-0.22	0.001
	Creatine vs. placebo	Week 1	-0.02	0.81	-0.08	0.24
		Week 6	0.02	0.81	-0.06	0.33
		Week 12	0.05	0.45	-0.1	0.07
	Creatine+400FA vs. placebo	Week 1	-0.12	0.13	-0.03	0.62
		Week 6	-0.1	0.1	-0.28	<0.001
		Week 12	-0.17	0.053	-0.14	0.01
%MMAs	FA400 vs. placebo	1 week	-0.85	0.07	-0.61	0.19
		6 weeks	-2.17	<0.001	-0.67	0.28
		12 weeks	-3.13	<0.001	-1.03	0.041
	FA800 vs. placebo	1 week	-0.87	0.09	-0.8	0.11
		6 weeks	-3.32	<0.001	-1.49	0.008
		12 weeks	-3.44	<0.001	-1.62	<0.001
	Creatine vs. placebo	1 week	-1.2	0.028	-0.36	0.49
		6 weeks	-2.07	0.003	-0.26	0.7
		12 weeks	-1.58	0.02	0.23	0.69
	Creatine+400FA vs. placebo	1 week	-1.24	0.007	-0.42	0.44
		6 weeks	-2.32	<0.001	-1.55	0.014
		12 weeks	-2.94	<0.001	-1.4	0.016

Arsenic metabolite	Treatment group comparison	Week	Low strata (≤ median) Mean difference	*P*	High strata (> median) Mean difference	*P*
%DMAs	FA400 vs. placebo	Week 1	1.44	0.26	1.09	0.38
		Week 6	3.25	0.005	3.22	0.02
		Week 12	4.79	<0.001	2.53	0.01
	FA800 vs. placebo	Week 1	2.57	0.07	2.4	0.054
		Week 6	4.08	<0.001	5.31	<0.001
		Week 12	5.71	<0.001	4.24	<0.001
	Creatine vs. placebo	Week 1	1.68	0.25	1.6	0.18
		Week 6	0.85	0.58	1.17	0.39
		Week 12	0.85	0.55	1.23	0.19
	Creatine+400FA vs. placebo	Week 1	2.8	0.035	0.67	0.59
		Week 6	3.68	0.001	4.94	<0.001
		Week 12	4.14	0.008	3.05	0.001

1. Placebo used as references. Models were adjusted for baseline As metabolite proportions (ln(%InAs), %MMAs, and %DMAs, respectively). Baseline choline median = 11.42 nmol/mL.
Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group; InAs, inorganic arsenic; MMAs, monomethyl-arsenical species; DMAs; dimethyl-arsenical species.

Table 4.6: Treatment effects[1] from linear models for the change over 12 weeks in ln(homocysteine), ln(GAA), ln(bAs), ln(%InAs), %MMAs, %DMAs, PMI, and ln(SMI) stratified by baseline choline, betaine, vitamin B_{12}, and folate.

			Low strata (≤ median)		**High strata (> median)**		**Wald test**	
Strata	**Change**	**Treatment**	***B* (95% CI)**	***P***	***B* (95% CI)**	***P***	***P***	***FDR***
Choline[2]	**ln(homocysteine)**	400FA	-0.26 (-0.35, -0.17)	<0.001	-0.25 (-0.35, -0.15)	<0.001	0.89	0.89
		800FA	-0.23 (-0.31, -0.15)	<0.001	-0.30 (-0.41, -0.18)	<0.001	0.39	0.78
		Creatine	-0.05 (-0.14, 0.05)	0.36	-0.02 (-0.12, 0.07)	0.6	0.76	0.89
		Creatine+400FA	-0.18 (-0.29, -0.08)	<0.001	-0.26 (-0.36, -0.16)	<0.001	0.29	0.78
	ln(GAA)[3]	400FA	0.02 (-0.07, 0.11)	0.68	0.04 (-0.05, 0.13)	0.4	0.77	0.77
		Creatine	-0.08 (-0.18, 0.03)	0.15	-0.21 (-0.32, -0.10)	<0.001	0.083	0.33
		Creatine+400FA	-0.09 (-0.19, 0.01)	0.07	-0.17 (-0.28, -0.06)	0.003	0.31	0.63
	ln(bAs)	400FA	0.00 (-0.15, 0.15)	0.96	0.14 (-0.03, 0.31)	0.1	0.21	0.44
		800FA	-0.10 (-0.28, 0.07)	0.24	-0.09 (-0.27, 0.09)	0.32	0.93	0.93
		Creatine	0.11 (-0.08, 0.3)	0.25	-0.04 (-0.20, 0.11)	0.6	0.22	0.44
		Creatine+400FA	0.00 (-0.19, 0.18)	0.99	-0.10 (-0.28, 0.08)	0.27	0.45	0.59
	ln(%InAs)[4]	400FA	-0.12 (-0.24, 0.01)	0.078	-0.12 (-0.24, 0.00)	0.048	0.955	0.955
		800FA	-0.15 (-0.27, -0.03)	0.017	-0.23 (-0.34, -0.12)	<0.001	0.344	0.688
		Creatine	0.05 (-0.09, 0.19)	0.466	-0.11 (-0.22, 0.00)	0.053	0.077	0.306
		Creatine+400FA	-0.12 (-0.30, 0.06)	0.185	-0.14 (-0.25, -0.03)	0.012	0.855	0.955
	%MMAs[5]	400FA	-3.00 (-4.09, -1.91)	<0.001	-1.25 (-2.25, -0.25)	0.014	0.02	0.078
		800FA	-3.23 (-4.34, -2.12)	<0.001	-1.76 (-2.72, -0.80)	<0.001	0.049	0.097
		Creatine	-1.47 (-2.88, -0.07)	0.04	0.05 (-1.09, 1.18)	0.936	0.098	0.108
		Creatine+400FA	-2.91 (-4.11, -1.72)	<0.001	-1.55 (-2.72, -0.38)	0.01	0.108	0.108
	%DMAs[6]	400FA	4.53 (2.03, 7.04)	<0.001	2.71 (0.71, 4.72)	0.008	0.265	0.931
		800FA	5.54 (3.21, 7.88)	<0.001	4.58 (2.72, 6.44)	<0.001	0.524	0.931
		Creatine	0.93 (-1.98, 3.83)	0.53	1.53 (-0.41, 3.47)	0.121	0.734	0.931
		Creatine+400FA	3.10 (-0.24, 6.44)	0.069	3.27 (1.35, 5.19)	0.001	0.931	0.931
	PMI[7]	400FA	-0.14 (-0.26, -0.02)	0.023	0.01 (-0.13, 0.15)	0.905	0.113	0.151
		800FA	-0.13 (-0.27, 0.00)	0.045	0.08 (-0.06, 0.21)	0.284	0.03	0.06
		Creatine	-0.16 (-0.3, -0.01)	0.035	0.06 (-0.07, 0.19)	0.369	0.028	0.06
		Creatine+400FA	-0.05 (-0.22, 0.13)	0.6	-0.01 (-0.15, 0.12)	0.855	0.762	0.762

Strata	Change	Treatment	Low strata (≤ median) *B* (95% CI)	*P*	High strata (> median) *B* (95% CI)	*P*	Wald test *P*	*FDR*
	ln(SMI)[8]	400FA	0.34 (0.22, 0.47)	<0.001	0.18 (0.07, 0.28)	0.001	0.047	0.181
		800FA	0.37 (0.25, 0.48)	<0.001	0.24 (0.14, 0.34)	<0.001	0.09	0.181
		Creatine	0.14 (0.01, 0.28)	0.035	0.02 (-0.09, 0.13)	0.691	0.172	0.229
		Creatine+400FA	0.30 (0.18, 0.42)	<0.001	0.21 (0.09, 0.34)	0.001	0.303	0.303
Betaine[9]	**ln(homocysteine)**	400FA	-0.24 (-0.34, -0.14)	<0.001	-0.26 (-0.35, -0.18)	<0.001	0.75	0.82
		800FA	-0.26 (-0.36, -0.15)	<0.001	-0.27 (-0.37, -0.17)	<0.001	0.82	0.82
		Creatine	-0.05 (-0.14, 0.05)	0.36	-0.02 (-0.11, 0.07)	0.64	0.73	0.82
		Creatine+400FA	-0.17 (-0.27, -0.06)	0.001	-0.28 (-0.38, -0.18)	<0.001	0.1	0.4
	ln(GAA)[3]	400FA	0.01 (-0.09, 0.10)	0.89	0.05 (-0.04, 0.15)	0.27	0.49	0.66
		Creatine	-0.14 (-0.25, -0.03)	0.011	-0.16 (-0.27, -0.05)	0.005	0.8	0.93
		Creatine+400FA	-0.10 (-0.2, 0.00)	0.057	-0.17 (-0.28, -0.06)	0.003	0.33	0.93
	ln(bAs)	400FA	0.01 (-0.14, 0.17)	0.87	0.11 (-0.05, 0.28)	0.17	0.37	0.93
		800FA	-0.13 (-0.30, 0.04)	0.12	-0.06 (-0.24, 0.12)	0.5	0.57	0.93
		Creatine	0.05 (-0.13, 0.22)	0.61	0.01 (-0.16, 0.18)	0.89	0.78	0.93
		Creatine+400FA	-0.05 (-0.23, 0.14)	0.61	-0.06 (-0.24, 0.12)	0.52	0.93	0.93
	ln(%InAs)[4]	400FA	-0.19 (-0.33, -0.05)	0.009	-0.06 (-0.17, 0.05)	0.281	0.153	0.306
		800FA	-0.29 (-0.43, -0.16)	<0.001	-0.11 (-0.22, -0.01)	0.035	0.042	0.167
		Creatine	-0.02 (-0.15, 0.12)	0.787	-0.07 (-0.19, 0.05)	0.264	0.589	0.786
		Creatine+400FA	-0.15 (-0.32, 0.02)	0.084	-0.13 (-0.26, -0.01)	0.035	0.877	0.877
	%MMA[5]	400FA	-3.06 (-4.20, -1.92)	<0.001	-1.35 (-2.33, -0.38)	0.007	0.026	0.102
		800FA	-2.93 (-4.02, -1.84)	<0.001	-2.14 (-3.18, -1.11)	<0.001	0.302	0.403
		Creatine	-1.09 (-2.44, 0.26)	0.112	-0.44 (-1.65, 0.77)	0.476	0.48	0.48
		Creatine+400FA	-2.78 (-4.02, -1.55)	<0.001	-1.69 (-2.86, -0.51)	0.005	0.206	0.403
	%DMA[6]	400FA	6.02 (3.29, 8.75)	<0.001	1.73 (-0.16, 3.62)	0.073	0.011	0.044
		800FA	7.07 (4.62, 9.51)	<0.001	3.49 (1.63, 5.35)	<0.001	0.022	0.044
		Creatine	2.03 (-0.77, 4.83)	0.155	1.03 (-1.15, 3.21)	0.353	0.579	0.579
		Creatine+400FA	4.15 (0.91, 7.38)	0.012	2.68 (0.54, 4.81)	0.014	0.455	0.579

Strata	Change	Treatment	Low strata (≤ median) *B* (95% CI)	*P*	High strata (> median) *B* (95% CI)	*P*	Wald test *P*	*FDR*
	PMI[7]	400FA	-0.11 (-0.24, 0.03)	0.122	-0.02 (-0.15, 0.11)	0.745	0.362	0.767
		800FA	0.02 (-0.13, 0.17)	0.807	-0.05 (-0.18, 0.07)	0.387	0.462	0.767
		Creatine	-0.08 (-0.23, 0.07)	0.319	-0.02 (-0.15, 0.11)	0.747	0.575	0.767
		Creatine+400FA	-0.03 (-0.2, 0.14)	0.71	0 (-0.14, 0.13)	0.95	0.806	0.806
	ln(SMI)[8]	400FA	0.39 (0.25, 0.53)	0	0.15 (0.06, 0.24)	0.001	0.005	0.021
		800FA	0.39 (0.27, 0.50)	0	0.23 (0.13, 0.33)	0	0.055	0.094
		Creatine	0.14 (0.01, 0.28)	0.033	0.04 (-0.07, 0.16)	0.455	0.258	0.258
		Creatine+400FA	0.34 (0.21, 0.47)	0	0.18 (0.07, 0.3)	0.002	0.071	0.094
Vitamin B_{12}[10]	**ln(homocysteine)**	400FA	-0.28 (-0.37, -0.19)	<0.001	-0.23 (-0.32, -0.14)	<0.001	0.45	0.75
		800FA	-0.25 (-0.34, -0.15)	<0.001	-0.28 (-0.39, -0.18)	<0.001	0.61	0.75
		Creatine	0.00 (-0.10, 0.10)	0.93	-0.05 (-0.14, 0.03)	0.23	0.46	0.75
		Creatine+400FA	-0.23 (-0.33, -0.13)	<0.001	-0.21 (-0.31, -0.11)	<0.001	0.75	0.75
	ln(GAA)[3]	400FA	0.05 (-0.04, 0.14)	0.26	0.00 (-0.09, 0.10)	0.95	0.46	0.93
		Creatine	-0.15 (-0.27, -0.03)	0.014	-0.14 (-0.24, -0.05)	0.004	0.93	0.93
		Creatine+400FA	-0.19 (-0.29, -0.09)	<0.001	-0.06 (-0.18, 0.05)	0.27	0.11	0.21
	ln(bAs)	400FA	-0.01 (-0.16, 0.15)	0.92	0.13 (-0.03, 0.29)	0.1	0.21	0.54
		800FA	-0.12 (-0.29, 0.06)	0.19	-0.08 (-0.25, 0.1)	0.4	0.74	0.93
		Creatine	0.03 (-0.15, 0.22)	0.72	0.02 (-0.14, 0.18)	0.78	0.93	0.93
		Creatine+400FA	-0.12 (-0.31, 0.07)	0.21	0.03 (-0.15, 0.2)	0.78	0.27	0.54
	ln(%InAs)[4]	400FA	-0.12 (-0.25, 0.01)	0.07	-0.06 (-0.20, 0.08)	0.43	0.51	0.68
		800FA	-0.20 (-0.32, -0.07)	0.002	-0.12 (-0.26, 0.02)	0.09	0.4	0.68
		Creatine	0.04 (-0.09, 0.16)	0.57	-0.07 (-0.21, 0.08)	0.36	0.28	0.68
		Creatine+400FA	-0.07 (-0.2, 0.05)	0.24	-0.08 (-0.25, 0.1)	0.4	0.99	0.99
	%MMA[5]	400FA	-2.56 (-3.75, -1.37)	<0.001	-1.84 (-3.13, -0.54)	0.006	0.42	0.74
		800FA	-2.68 (-3.96, -1.40)	<0.001	-2.16 (-3.42, -0.89)	0.001	0.57	0.74
		Creatine	-1.34 (-2.96, 0.28)	0.11	-0.42 (-1.85, 1.01)	0.56	0.4	0.74
		Creatine+400FA	-2.14 (-3.48, -0.79)	0.002	-1.80 (-3.22, -0.38)	0.013	0.74	0.74

Strata	Change	Treatment	Low strata (≤ median) *B* (95% CI)	*P*	High strata (> median) *B* (95% CI)	*P*	Wald test *P*	*FDR*
	%DMA[6]	400FA	3.97 (1.63, 6.32)	0.001	2.90 (0.33, 5.46)	0.027	0.54	0.64
		800FA	5.17 (2.98, 7.36)	<0.001	4.23 (1.69, 6.77)	0.001	0.58	0.64
		Creatine	0.72 (-1.85, 3.29)	0.58	2.09 (-0.53, 4.71)	0.12	0.46	0.64
		Creatine+400FA	2.83 (0.54, 5.13)	0.02	1.88 (-1.45, 5.20)	0.27	0.64	0.64
	PMI[7]	400FA	-0.08 (-0.22, 0.06)	0.24	-0.11 (-0.27, 0.04)	0.15	0.764	0.764
		800FA	-0.01 (-0.15, 0.13)	0.9	-0.07 (-0.23, 0.09)	0.41	0.585	0.764
		Creatine	-0.14 (-0.31, 0.03)	0.1	-0.02 (-0.19, 0.15)	0.81	0.324	0.764
		Creatine+400FA	-0.09 (-0.25, 0.07)	0.26	-0.01 (-0.19, 0.17)	0.9	0.523	0.764
	ln(SMI)[8]	400FA	0.29 (0.17, 0.41)	<0.001	0.23 (0.09, 0.37)	0.001	0.483	0.782
		800FA	0.31 (0.19, 0.43)	<0.001	0.27 (0.14, 0.39)	0	0.62	0.782
		Creatine	0.12 (-0.02, 0.26)	0.1	0.08 (-0.06, 0.21)	0.28	0.671	0.782
		Creatine+400FA	0.25 (0.12, 0.37)	<0.001	0.22 (0.08, 0.35)	0.002	0.782	0.782
Folate[11]	**ln(homocysteine)**	400FA	-0.33 (-0.43, -0.23)	<0.001	-0.16 (-0.25, -0.08)	<0.001	0.011	0.024
		800FA	-0.34 (-0.45, -0.23)	<0.001	-0.20 (-0.29, -0.11)	<0.001	0.049	0.065
		Creatine	-0.05 (-0.15, 0.06)	0.38	-0.02 (-0.11, 0.07)	0.67	0.7	0.7
		Creatine+400FA	-0.31 (-0.42, -0.21)	<0.001	-0.14 (-0.22, -0.05)	0.002	0.012	0.024
	ln(GAA)[3,12]	400FA	0.02 (-0.06, 0.10)	0.558	-0.04 (-0.12, 0.05)	0.397	0.309	0.309
		Creatine	-0.10 (-0.20, 0.00)	0.052	-0.22 (-0.32, -0.13)	0	0.091	0.136
		Creatine+400FA	-0.09 (-0.18, 0.01)	0.078	-0.21 (-0.31, -0.11)	0	0.077	0.136
	ln(bAs)	400FA	0 (-0.14, 0.15)	0.97	0.14 (-0.03, 0.31)	0.2	0.25	0.49
		800FA	-0.14 (-0.31, 0.03)	0.1	-0.05 (-0.23, 0.13)	0.57	0.46	0.62
		Creatine	0.01 (-0.16, 0.19)	0.88	0.04 (-0.13, 0.21)	0.65	0.84	0.84
		Creatine+400FA	-0.15 (-0.34, 0.05)	0.13	0.04 (-0.13, 0.21)	0.66	0.16	0.49
	ln(%InAs)[4]	400FA	-0.13 (-0.24, -0.01)	0.03	-0.12 (-0.26, 0.02)	0.098	0.944	0.944
		800FA	-0.22 (-0.32, -0.12)	<0.001	-0.17 (-0.3, -0.03)	0.019	0.555	0.944
		Creatine	0.02 (-0.10, 0.14)	0.746	-0.10 (-0.23, 0.03)	0.145	0.195	0.78
		Creatine+400FA	-0.13 (-0.27, 0.02)	0.095	-0.15 (-0.3, 0.00)	0.055	0.818	0.944

Strata	Change	Treatment	Low strata (≤ median) *B* (95% CI)	*P*	High strata (> median) *B* (95% CI)	*P*	Wald test *P*	*FDR*
	%MMA[5]	400FA	-2.68 (-3.66, -1.69)	<0.001	-1.58 (-2.74, -0.43)	0.008	0.155	0.62
		800FA	-2.52 (-3.55, -1.49)	<0.001	-2.47 (-3.55, -1.39)	<0.001	0.944	0.944
		Creatine	-0.50 (-1.73, 0.72)	0.417	-0.79 (-2.08, 0.49)	0.226	0.749	0.944
		Creatine+400FA	-2.30 (-3.58, -1.02)	0.001	-2.1 (-3.23, -0.97)	<0.001	0.818	0.944
	%DMA[6]	400FA	4.09 (1.98, 6.20)	<0.001	3.26 (0.65, 5.87)	0.015	0.626	0.859
		800FA	5.26 (3.53, 6.99)	<0.001	4.97 (2.36, 7.58)	<0.001	0.859	0.859
		Creatine	0.27 (-2.14, 2.67)	0.827	2.32 (-0.31, 4.96)	0.083	0.256	0.859
		Creatine+400FA	2.96 (0.11, 5.81)	0.042	3.67 (0.95, 6.39)	0.008	0.722	0.859
	PMI[7]	400FA	-0.09 (-0.21, 0.04)	0.175	-0.02 (-0.16, 0.11)	0.764	0.483	0.679
		800FA	0.01 (-0.13, 0.15)	0.889	-0.05 (-0.18, 0.08)	0.411	0.509	0.679
		Creatine	-0.10 (-0.24, 0.04)	0.173	0.02 (-0.11, 0.16)	0.713	0.209	0.679
		Creatine+400FA	-0.02 (-0.18, 0.13)	0.77	-0.02 (-0.17, 0.13)	0.839	0.943	0.943
	ln(SMI)[8]	400FA	0.32 (0.21, 0.43)	<0.001	0.20 (0.07, 0.32)	0.003	0.144	0.575
		800FA	0.31 (0.20, 0.41)	<0.001	0.30 (0.18, 0.41)	<0.001	0.894	0.894
		Creatine	0.05 (-0.07, 0.17)	0.402	0.11 (-0.02, 0.23)	0.089	0.51	0.894
		Creatine+400FA	0.27 (0.14, 0.4)	<0.001	0.24 (0.12, 0.36)	<0.001	0.729	0.894

1. Placebo used as reference group. 2. Baseline choline median = 11.42 nmol/mL. 3. GAA was measured in a subset of participants; 400FA and 800FA groups were excluded from analyses. 4. Adjusted for baseline ln(%InAs). 5. Adjusted for baseline %MMAs. 6. Adjusted for baseline %DMAs. 7. Adjusted for baseline ln(PMI). 8. Adjusted for baseline ln(SMI). 9. Baseline betaine median = 43.63 nmol/mL. 10. Baseline vitamin B_{12} median = 214.89 pmol/L. 11. Baseline plasma folate median = 13.50 nmol/L. 12. Adjusted for baseline ln(GAA).

Abbreviations: 400FA, 400 μg FA/day treatment group; 800FA, 800 μg FA/day treatment group; Creatine, 3 g creatine/day treatment group; creatine+400FA, 3 g creatine and 400 μg FA/day treatment group; bAs, blood arsenic; GAA, guanidinoacetate; InAs, inorganic arsenic; MMAs, monomethyl-arsenical species; DMAs; dimethyl-arsenical species; PMI, primary methylation index; SMI, secondary methylation index.

Chapter 5

Locus-specific differential DNA methylation and urinary arsenic: an epigenome-wide association study among adults with low-to-moderate arsenic exposure

Anne K. Bozack,[1] Maria Tellez-Plaza,[2,3] Karin Haack,[4] Mary Gamble,[1] Arce Domingo,[1] Jason G. Umans,[5,6] Lyle G. Best,[7] Joseph Yracheta,[8] Matthew O Gribble,[9] Andres Cardenas,[10] Kevin A Francesconi,[11] Walter Goessler,[11] Wan-Yee Tang,[3] M Daniele Fallin,[12,13] Shelley A Cole,[1] Ana Navas-Acien[1]

Affiliations: 1. Dept. of Environmental Health Science, Columbia University; 2. Biomedical Research Institute INCLIVA, Hospital Clinico Universitario de Valencia, University of Valencia, Valencia, Spain; 3. Dept. of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health; 4. Dept. of Genetics, Texas Biomedical Research Institute; 5. MedStar Health Research Institute; 6. Georgetown/Howard Universities Center for Clinical and Translational Sciences; 7. Missouri Breaks Industries Research; 8. Dept. of Pharmaceutics, University of

Washington; 9. Department of Environmental Health, Emory University; 10. Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley; 11. Institute of Chemistry, University of Graz, Graz, Austria; 12. Dept. of Mental Health, Johns Hopkins University Bloomberg School of Public Health; 13. Wendy Klag Center for Autism and Developmental Disabilities.

Submitted for tribal review.

5.1 Abstract

Background: Chronic exposure to arsenic (As), a human toxicant and carcinogen, remains a global public health problem. Health risks persist after As exposure has ended, suggesting epigenetic dysregulation as a mechanistic link between exposure and health outcomes.

Objectives: We investigated the association between total urinary As and locus-specific DNA methylation among American Indian adults with low-to-moderate As exposure.

Methods: DNA methylation was measured in 2,325 participants using the Illumina MethylationEPIC array. We implemented linear regression models to test differentially methylated positions (DMPs) and used *DMRcate* to identify regions (DMRs). We also conducted gene ontology enrichment analysis. Models were adjusted for estimated cell type proportions, age, sex, BMI, smoking, education, estimated glomerular filtration rate, and study center. Arsenic was measured in urine as the sum of inorganic and methylated species.

Results: In fully adjusted models, methylation at 20 individual CpGs was associated with urinary As after false discovery rate (FDR) correction ($FDR < 0.05$). After Bonferroni correction, 5 CpGs remained associated with total urinary As levels ($P_{Bonferroni} < 0.05$), located in *SLC7A11*, *ANKS3*, *LINGO3*, *CSNK1D*, *ADAMTSL4*. We identified one DMR on chromosome 11 (chr11:2,322,050-2,323,247), annotated to *C11orf2*; *TSPAN32* genes.

Conclusion: This is one of the first epigenome-wide association studies to investigate the association between As exposure and locus-specific DNA methylation using the Illumina MethylationEPIC array and the largest epigenome wide studies of arsenic exposure. The top DMP was located in *SLC7A11A*, a gene involved in cystine/glutamate transport and the biosynthesis of glutathione, an antioxidant that

may protect against As-induced oxidative stress. Additional DMPs were located in genes associated with tumor development and glucose metabolism. Further research is needed to investigate whether As-related DNA methylation signatures are biomarkers of disease development.

Acknowledgments

The authors would like to thank all Strong Heart Study participants and staff who have made this work possible.

5.2 Background

Arsenic (As) exposure through drinking water is a global public health concern affecting at least 140 million people (World Health Organization, 2012). In 2001, the United States (US) Environmental Protection Agency (EPA) lowered maximum contaminant level (MCL) from 50 to 10 μg/L. With the implementation of this new standard, the number of people served by public water systems with water arsenic > 10 μg/L has declined from 13 million in 2001 to 296,000 in 2018 U.S. Environmental Protection Agency (2018). An estimated 2.1 million individuals, however, remain exposed to As in drinking water from private wells (Ayotte et al.), which are not regulated under the MCL.

Arsenic is a human toxicant and group 1 carcinogen (World Health Organization, 2011). Chronic As exposure increases the risk of numerous health conditions including skin lesions, impaired intellectual function, cardiovascular disease, diabetes, inflammation, and cancers including bladder, lung, kidney, liver, skin, and possibly prostate (IARC Working Group, 2009; National Research Council, 2013; Moon et al., 2017). Elevated risk of cancer mortality (Roh et al., 2018; Smith et al., 2018) and lung disease (Steinmaus et al., 2016) following early-life exposure persists decades after exposure has been reduced. Epigenetic dysregulation, including changes in DNA methylation patterns, may provide a mechanistic link between As exposure and health outcomes with prolonged latency periods (Bailey et al., 2016). DNA methylation can influence gene expression by inhibiting transcription factor binding in promoter regions and recruiting DNA binding proteins, and is involved in maintaining chromosomal stability (Robertson, 2005).

In vitro, animal, and human evidence supports that alterations in the epigenome are involved in the etiology of As-induced health outcomes and carcinogenesis (Bailey et al., 2016; Bailey and Fry, 2014; Carlin et al., 2015). In epidemiological studies, As ex-

posure has been associated with global DNA methylation levels (Intarasunanont et al., 2012; Kile et al., 2012; Niedzwiecki et al., 2013, 2015; Pilsner et al., 2009, 2007). Locus-specific DNA methylation has also been assessed in epigenomic-wide association studies (EWAS) using the Illumina Infinium HumanMethylation BeadChip (450K), which interrogates DNA methylation at > 480,000 loci (Argos, 2015). Studies of in utero As exposure using cord blood (Broberg et al., 2014; Kaushal et al., 2017; Kile et al., 2014; Rojas et al., 2015) and studies of adults (Ameer et al., 2017; Argos et al., 2015; Demanelis et al., 2019) have identified significant associations with DNA methylation at individual CpG sites; however the number of CpGs identified and their identity differ between studies.

Inconsistent results between epidemiological studies of the association between As exposure and epigenetic dysregulation may be due to differences between populations studied (e.g., age, sex, genetic structure), levels of As exposure, differential residual confounding, and methods for quantifying DNA methylation. In addition, EWAS have had limited statistical power due to small sample sizes, particularly when interrogating a large number of CpG sites (Argos, 2015). The objective of this study was to investigate the association between total urinary As and locus-specific DNA methylation in the Strong Heart Study (SHS), a population-based prospective cohort of American Indian adults with low-to-moderate levels of As exposure primarily through drinking water (Navas-Acien et al., 2009). DNA methylation in the SHS has been measured in 2,325 study participants using the Infinium MethylationEPIC BeadChip, which covers > 850,000 CpG sites.

5.3 Methods

Study population

The SHS has been described in detail (Lee et al., 1990). Briefly, from 1989-1991, SHS participants were recruited from 13 tribes in Arizona, Oklahoma, and North and South Dakota. All tribal members (a random-stratified subset in North and South Dakota) aged 45-74 years were eligible for enrollment. At baseline, 4,549 adults were enrolled with a participation rate of 62%. Most participants were born in their communities and have lived there for their entire lives.

Ethics

The study protocols were approved by Institutional Review Boards from the participating research institutions, the Indian Health Service, and the tribal communities. One community withdrew consent for future studies, resulting in a sample size 3,516. Participating communities approved this manuscript and received a lay summary of findings.

Data collection

During baseline visit, a clinical exam was conducted and bio-specimens were collected by trained and certified nurses and medical examiners. Urinary As concentration: Analytical methods and quality control for urine As measurement has been described in detail (Scheer et al., 2012). In summary, baseline spot urine samples were stored in polypropylene tubes and frozen samples were shipped on dry ice to the Penn Medical Laboratory, MedStar Research Institute (Washington, DC, USA), where they were stored at $<$ -70°C. In 2009-2010, samples were thawed and an aliquot up to 1.0 mL was transported on dry ice to the Trace Element Laboratory, Graz University (Austria),

where they were stored at < -70°C until analysis for total urinary As concentrations and As metabolite concentrations.

Urinary As concentrations were calculated as the sum of the concentrations of inorganic (InAs), monomethyl (MMAs), and dimethyl (DMAs) arsenic species, and are referred to in the manuscript as total urinary As. InAs, MMAs, and DMAs concentrations were measured using high-performance liquid chromatography coupled to inductively coupled plasma mass spectrometry (Agilent 1100 HPLC and Agilent 7700x ICP-MS, Agilent Technologies, Santa Clara, California) (Scheer et al., 2012). Inter-assay coefficients of variation: 6.0%, 6.5%, 5.9% for InAs, MMAs, and DMAs, respectively. The limits of detection (LOD) for InAs (As^{III} + As^{V}), MMAs, and DMAs was 0.1 μg/L (Scheer et al., 2012). Samples with arsenic species concentrations beneath the LOD were replaced with LOD/$\sqrt{2}$ (InAs: N = 128, 5.5%; MMAs: N = 17, 0.7%; DMAs: N = 0). Concentrations of arsenobetaine, a nontoxic arsenic species found in seafood (Joint FAO/WHO Export Committee on Food Additives, 1989), were low (median 0.65 μg/g creatinine), reflecting little seafood intake in the study population.

Epigenome-wide DNA methylation assessment and quality control

Blood samples were collected in EDTA tubes and DNA from white blood cells was isolated and stored at <-70°C at the Penn Medical Laboratory, MedStar Health Research Institute. For epigenetic analyses, DNA samples were shipped to the Texas Biomedical Research Institute. Genomic DNA was bisulfite-converted and eluted in buffer. After excluding participants with cardiovascular disease at baseline, missing data for urinary metal concentrations and/or baseline sociodemographics or health status, and insufficient DNA (Appendix Table 5.5), DNA methylation was measured in 2,352 samples (Appendix Figure 5.4). Participants included in analyses of DNA

methylation were similar to all eligible participants (Appendix Table 5.5). DNA methylation was measured using the Infinium MethylationEPIC BeadChip (850K) (Illumina, San Diego, CA) according to the manufacturers instructions with samples randomized within and across plates to minimize potential batch artifacts. Replicate and across-plate samples were included on each plate. The MethylationEPIC BeadChip provides a measure of DNA methylation at a single nucleotide resolution at $>$ 850,000 methylation sites, including $>$ 90% of loci measured by the 450K microarray (Illumina, 2015).

Raw methylation image files were processed using the minfi package (Aryee et al., 2014) in R (R Core Team, 2015). Density plots were generated to analyze the distribution of beta values, and 18 samples that did not have classical bimodal distributions were excluded. Normalization was performed using single sample Noob (Fortin et al., 2017). An additional eight samples were excluded with low median intensity of methylated and unmethylated channels ($\log_2$(intensity) $<$ 10). Probes determined to be technical failures (P-detection $>$ 0.01 in $\geq$ 5% of samples) were removed. Batch effects for plate and row were corrected using the combat function in the sva package, which employs an empirical Bayesian framework (Leek et al., 2012). Clustering by batch and row before and after applying combat was visually assessed using the first two principal components. The proportions of CD8+ T cells, CD4+ T cells, natural killer (NK) cells, B cells, monocytes, and granulocytes in each sample was estimated using the Houseman projection method (Houseman et al., 2012). In addition, probes located in X and Y chromosomes, probes associated with SNPs with a minor allele frequency $>$ 5% in an admixed American population, and probes previously identified as cross-reactive were removed prior to analysis (McCartney et al., 2016; Pidsley et al., 2016). After exclusion of probes and samples, DNA methylation data measured at 788,753 loci in 2,325 samples were available for analysis. To validate our data processing and quality control process, we tested for epigenome-wide associations

with smoking. Classical smoking-induced differentially methylated positions (DMPs) were highly significant in our study (Domingo et al., 2019, under review).

Other variables

Height and weight were measured for BMI calculation. A trained interviewer collected data on sociodemographics and health-related behaviors (e.g., age, sex, history of smoking) (Lee et al., 1990). Diabetes status was determined according to the American Diabetes Association classification using fasting glucose $\geq$ 126 mg/dL, 2-h post-load plasma glucose $\geq$ 200 mg/dL, hemoglobin HbA1c $\geq$ 6.5%, or taking diabetes medication (American Diabetes Association, 2014).

Urinary creatinine was measured at the National Institute of Diabetes and Digestive and Kidney Diseases Epidemiology and Clinical Research Branch laboratory (Phoenix, AZ, USA) using an automated alkaline picrate methodology run on a rapid flow analyzer (Lee et al., 1990). Total urinary As concentrations (μg/L) were divided by urinary creatinine (g/L) to account for urine dilution. Serum creatinine was measured in fasting blood samples on a Hitachi 717 platform (Hitachi Ltd.) using an automated alkaline-picrate rate method (Roche Diagnostics). Serum creatinine, age, and sex were used to calculate estimated glomerular filtration rate (eGFR) using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation (Levey et al., 2009; Shara et al., 2012).

Statistical analysis

Descriptive statistics (means and SDs for continuous variables, frequencies for categorical variables) were calculated for all covariates. To adjust for between-individual differences in urine dilution, total urinary As concentrations were divided by urinary creatinine concentrations and expressed as μg/g creatinine. Adjusted urinary

As levels were right skewed and natural log-transformed to reduce the influence of extreme values. Associations between estimated cell types proportions and ln(total urinary As) were evaluated using linear models adjusted for age, sex BMI and smoking status (never smoker, former smoker, current smoker).

DMPs were identified using linear regression models implemented in the *limma* package in R with empirical Bayes smoothing of standard errors (Ritchie et al., 2015). Beta-values were logit transformed to M-values (i.e., M-value = ln(Beta-value/(1-Beta-value))) to meet linear regression model assumptions (Du et al., 2010). Models were adjusted for sex, age, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), eGFR, and cell type proportion estimates. Potential systematic biases were evaluated using QQ-plots and the genomic inflation factor (λ). For the fully adjusted models, $\lambda = 0.923$, suggesting that our analyses were not impacted by genomic inflation (Appendix Figure 5.5). Multiple comparisons were accounted for using the Benjamini and Hochberg method for false discovery rates (FDR) (Benjamini and Hochberg, 1995) and the Bonferroni correction. Using the *p.adjust* function in R, FDR-adjusted P-values were calculated as the number of expected P-values $\leq P$ divided by the number of observed P-values $\leq P$ (i.e., for N ordered P-values P_1 through P_N, $FDR = (P_i \text{ x N})/i$); Bonferroni-adjusted P-values were calculated as P multiplied by the number of tests (i.e., $P_{i,Bonferroni} = P_1$ x N).

Sensitivity analyses were performed for the potential effect modifiers of sex, study center, and diabetes status. For participants without diabetes, specific gravity was used to correct total urinary As concentrations (μg/L) for between-individual differences in urine dilution by multiplying the sample urinary As concentration by the ratio (mean specific gravity - 1)/(participant's specific gravity - 1) (Miller et al., 2004). *limma* analyses were performed for each strata and nominal P-values were calculated.

Differentially methylation regions (DMRs) were tested using the *DMRcate* package in R, which uses a Gaussian kernel smoothing function to grouping of significant probes identified by *limma* (Peters et al., 2015). *DMRcate* was applied using a Gaussian kernel bandwidth λ = 1,000 with a smoothing factor C = 2. The bandwidth value defines the maximum distance in nucleotides used to group methylated loci.

Gene ontology (GO) analysis was conducted using the *GOmeth* (Geeleher et al., 2013) function implemented in the R package *missMethyl* (Phipson et al., 2015). The *GOmeth* algorithm identifies GO terms that contain an overrepresentation of genes with DMPs while accounting for differences in the *a priori* probabilities of genes to include DMPs based on representation among probes included on the MethylationEPIC BeadChip. All probes tested in limma analyses and probes identified as significantly associated with urinary As levels at $FDR < 0.05$ were used as input for the *GOmeth* function.

Evaluation of previously identified signals

To identify previous studies investigating the association between As exposure and loci-specific DNA methylation, we searched PubMed for the terms arsenic and DNA methylation. Loci associated with As exposure reported by those studies measuring DNA methylation using the Illumina HumanMethylation BeadChip or the Infinium MethylationEPIC BeadChip were included in a lookup approach. Loci with a nominal $P < 0.05$ in the *limma* analysis described above were considered statistically significant.

All analyses were performed using R 3.4.3 (R Core Team, 2015).

5.4 Results

Participant characteristics

The mean (SD) age of participants was 56.2 (8.1) years, 58.5% of participants were female, and 41.6% of participants had diabetes (Table 5.1). Total urinary As levels ranged from 1.7 - 113.0 μg/g creatinine, with mean (SD) 11.7 (10.6) μg/g creatinine and median (interquartile range, IQR) 8.6 (5.2, 14.4) μg/g creatinine. Total urinary As was associated with the imputed proportions of NK cells and B cells: on average, with every one-unit increase in ln(μg total As/g creatinine), there was an increase of 5% in the proportion of NK cells ($P = 0.004$), and a decrease of 5% in the proportion of B cells ($P < 0.001$) (Appendix Table 5.6). Total urinary As was borderline associated with the imputed proportions of monocytes ($P = 0.06$) and not associated with the imputed proportions of CD8+ T cells, CD4+ T cells, and granulocytes.

Differentially methylated positions

In locus-specific analyses, 788,753 methylated positions were tested for associations with ln(total urinary As), of which 39,857 (5.1%) were significantly associated with As at a nominal $P < 0.05$. After adjusting for multiple comparisons using an FDR and Bonferroni approach, 20 (FDR < 0.05) and five ($P_{Bonferroni} < 0.05$) loci remained significantly associated with ln(total urinary As), respectively (Figure 5.1). Table 5.2 summarizes the loci significant at the FDR threshold, and includes effect size estimates from models of Beta-values for interpretation of results. Most significant loci (N = 13) were located within gene bodies. Two genes, leucine rich repeat and Ig domain containing 3 (*LINGO3*) and casein kinase 1 delta (*CSNK1D*) contained two FDR-significant CpGs located within the same genomic feature (in *LINGO3*, cg22294740 and cg08059112 are located 74 nucleotides apart; in *CSNK1D*, cg20493718 and cg21369801

Table 5.1: Participant characteristics (N = 2,325).

	N	(%)
Female	1,361	(58.5)
Age, mean years (SD)	56.2	(8.1)
Total urinary As, mean μg/g creatinine (SD)	11.7	(10.6)
< 5 μg/g creatinine	533	(22.9)
5 - 10 μg/g creatinine	1,252	(53.8)
> 10 μg/g creatinine	540	(23.2)
Study center		
Arizona	312	(13.4)
Oklahoma	981	(42.2)
North and South Dakota	1,032	(44.4)
Education		
< High school diploma	963	(41.4)
High school diploma or GED	658	(28.3)
> High school diploma	704	(30.3)
Smoking		
Never smoker	684	(29.4)
Former smoker	748	(32.2)
Current smoker	893	(38.4)
BMI, mean (SD)	30.3	(6.1)
Diabetes[a]	968	(41.6)

a. Fasting glucose (≥ 126 mg/dL), 2-h post-load plasma glucose (≥ 200 mg/dL), HbA1c (≥ 6.5%), or taking diabetes medication.

are located 18 nucleotides apart). Of the 20 FDR-significant loci, the association between urinary As levels and DNA methylation was positive at 18 loci and inverse at two loci (cg06690548 and cg00500428). Among all probes, urinary As levels also appeared to be related to hypermethylation: 87% of the top 100 probes ranked by P-values were positively associated with As, and 58% of all probes were positively associated with As (Figure 5.2).

Sensitivity analyses

Results were consistent in stratified analyses by sex (Appendix Table 5.7), study center (Appendix Table 5.8), and diabetes status (Appendix Table 5.9). By sex,

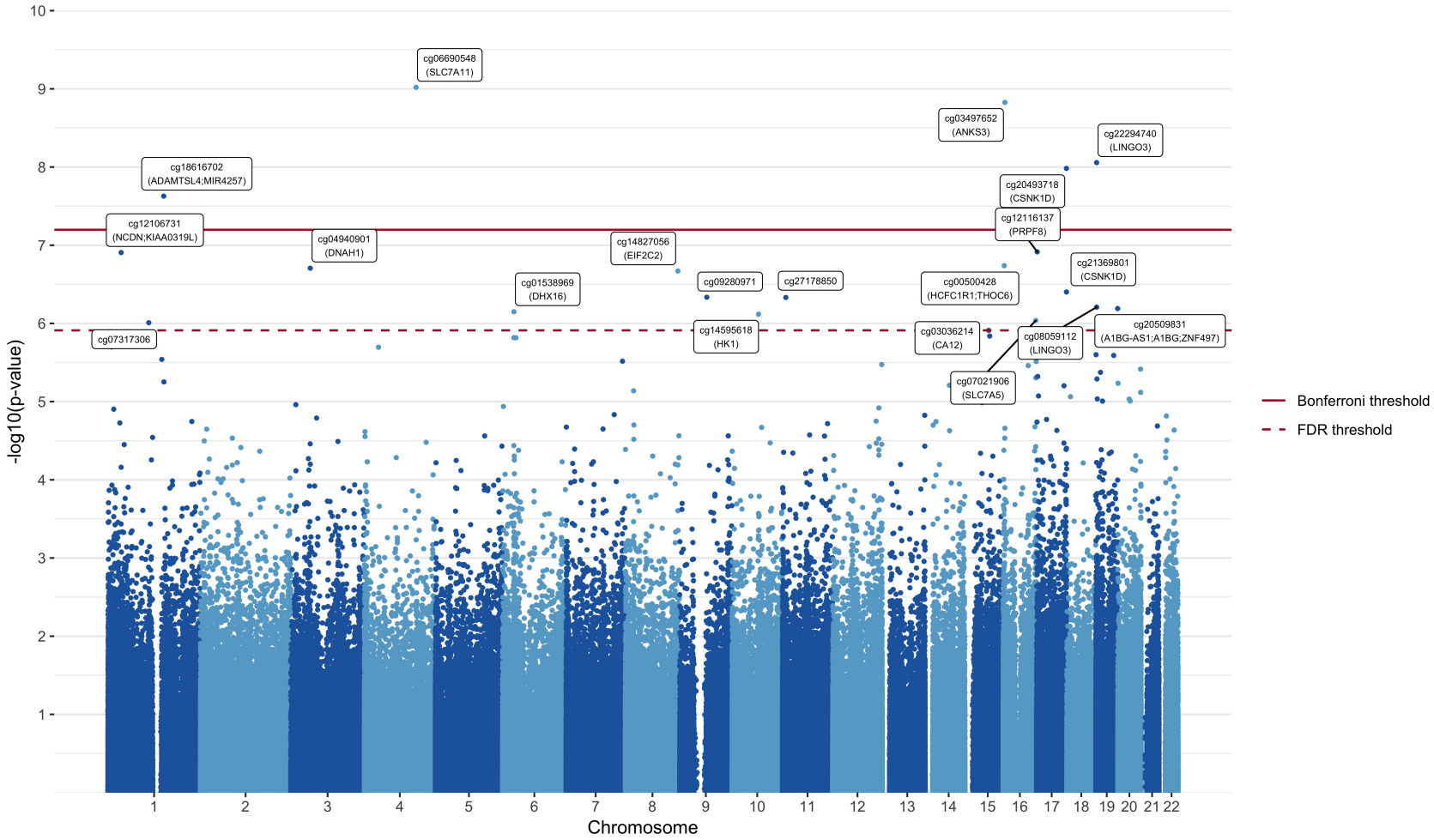


Figure 5.1: Manhattan plot for the epigenome-wide association of total urinary arsenic levels and methylation levels. *limma* models adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. The solid line represents the Bonferroni threshold and the dashed line represents the FDR threshold for significance.

Table 5.2: Differentially methylated positions associated with total urinary arsenic levels at $FDR < 0.05$, sorted by chromosome and position.

CpG	Chr	Position[a]	Gene	Gene region	Odds of DNAm (95% CI)[b]	*P*	*FDR*	$P_{Bonferroni}$	Median % DNAm[c]	Mean change in % DNAm (95% CI)[d]
cg12106731	1	36023279	NCDN; KIAA0319L	TSS200; TSS1500	1.05 (1.03, 1.07)	1.24×10^{-7}	0.014	0.098	0.60	0.04 (0.02, 0.05)
cg07317306	1	110314824	Intergenic		1.03 (1.02, 1.05)	9.79×10^{-7}	0.041	0.773	17.75	0.50 (0.31, 0.69)
cg18616702	1	150523808	ADAMTSL4; MIR4257	5'UTR; TSS1500	1.05 (1.03, 1.07)	2.35×10^{-8}	0.004	0.019	22.45	0.80 (0.50, 1.10)
cg04940901	3	52418250	DNAH1	Body	1.05 (1.03, 1.07)	1.96×10^{-7}	0.017	0.155	71.58	0.97 (0.60, 1.33)
cg06690548	4	139162808	SLC7A11	Body	0.91 (0.88, 0.94)	9.56×10^{-10}	0.001	0.001	96.24	-0.49 (-0.66, -0.32)
cg01538969	6	30624636	DHX16	Body	1.06 (1.04, 1.08)	7.09×10^{-7}	0.035	0.559	57.05	1.31 (0.80, 1.82)
cg14827056	8	141550539	EIF2C2	Body	1.07 (1.04, 1.10)	2.14×10^{-7}	0.017	0.169	88.76	0.58 (0.35, 0.81)
cg09280971	9	73036509	Intergenic		1.03 (1.02, 1.04)	4.63×10^{-7}	0.028	0.365	31.54	0.64 (0.40, 0.89)
cg14595618	10	71135446	HK1	Body	1.06 (1.04, 1.08)	7.63×10^{-7}	0.035	0.602	39.52	1.29 (0.76, 1.82)
cg27178850	11	9039431	Intergenic		1.04 (1.02, 1.06)	4.66×10^{-7}	0.028	0.367	69.43	0.82 (0.51, 1.13)
cg03036214	15	63640658	CA12	Body	1.06 (1.03, 1.08)	1.23×10^{-6}	0.048	0.967	56.18	1.30 (0.78, 1.82)
cg00500428	16	3074502	HCFC1R1; THOC6	TSS1500; Body	0.94 (0.92, 0.96)	1.83×10^{-7}	0.017	0.144	5.61	-0.31 (-0.43, -0.19)
cg03497652	16	4751569	ANKS3	Body	1.09 (1.06, 1.13)	1.49×10^{-9}	0.001	0.001	74.34	1.57 (1.08, 2.06)
cg07021906	16	87866833	SLC7A5	Body	1.07 (1.04, 1.09)	9.19×10^{-7}	0.04	0.725	82.02	0.85 (0.51, 1.20)
cg12116137	17	1576449	PRPF8	Body	1.09 (1.05, 1.12)	1.22×10^{-7}	0.014	0.096	77.19	1.33 (0.83, 1.84)
cg21369801	17	80202961	CSNK1D	Body, 3'UTR	1.08 (1.05, 1.11)	3.96×10^{-7}	0.028	0.312	77.81	1.17 (0.73, 1.62)
cg20493718	17	80202979	CSNK1D	Body, 3'UTR	1.13 (1.08, 1.17)	1.04×10^{-8}	0.002	0.008	97.47	0.35 (0.20, 0.50)
cg08059112	19	2294887	LINGO3	5'UTR	1.04 (1.02, 1.05)	6.19×10^{-7}	0.034	0.488	56.18	0.89 (0.54, 1.24)
cg22294740	19	2294961	LINGO3	5'UTR	1.09 (1.06, 1.12)	8.78×10^{-9}	0.002	0.007	55.92	1.84 (1.20, 2.48)
cg20509831	19	58866362	A1BG-AS1; A1BG; ZNF497	Body; TSS1500; 3'UTR	1.05 (1.03, 1.07)	6.45×10^{-7}	0.034	0.509	28.1	0.89 (0.53, 1.25)

a. GRCh37/hg19 assembly. b. Effect size estimate from limma models of M-values (i.e., odds of being methylated to being unmethylated for each unit change in ln(total urinary arsenic levels)) adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. c. Percent methylation calculated Beta-value x 100. d. Effect size estimate from limma models of Beta-values (i.e., mean change in percent methylation for each unit change in ln(total urinary arsenic levels)) adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. DNAm, DNA methylation.

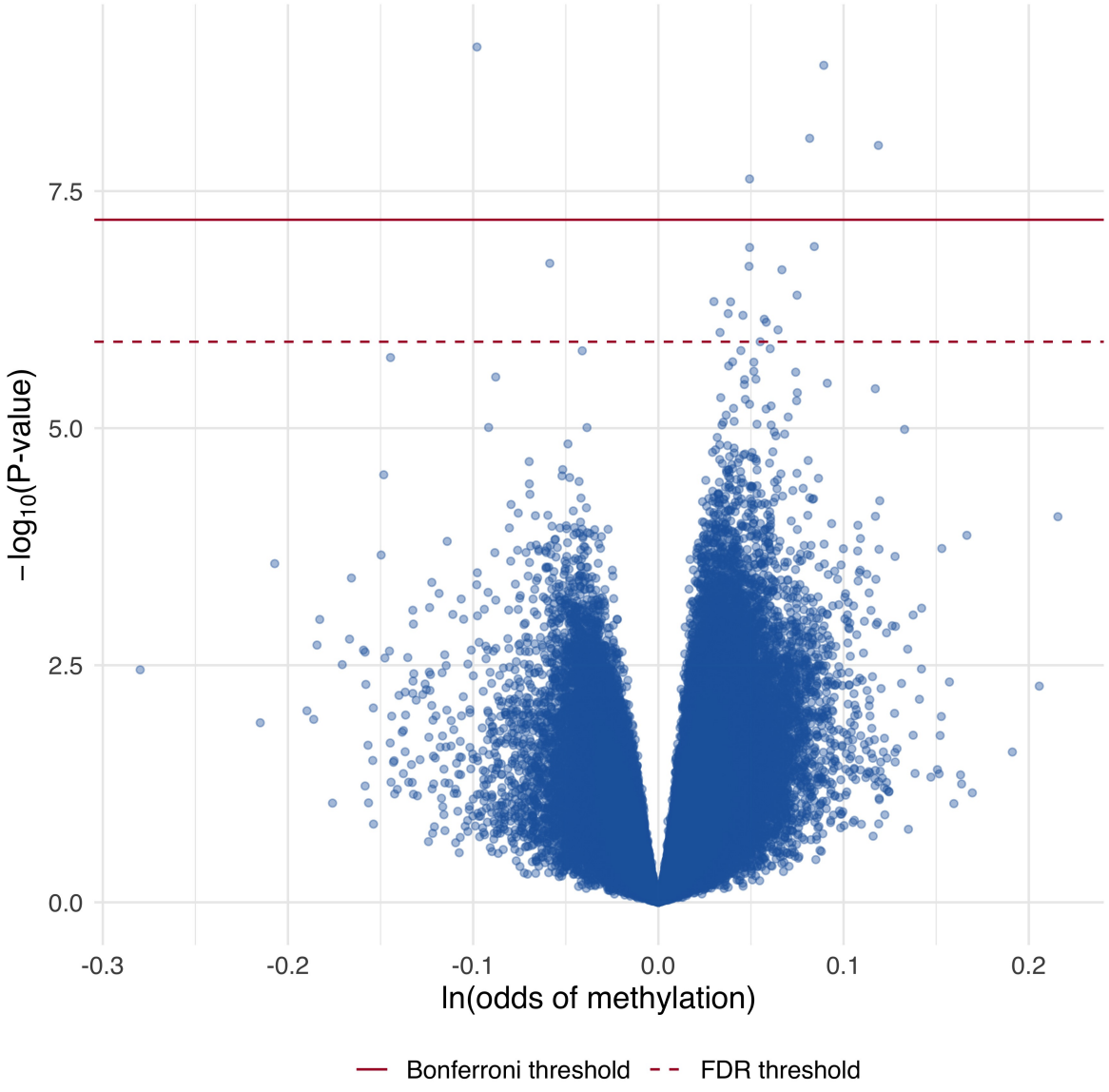


Figure 5.2: Volcano plot for the epigenome-wide association of total urinary arsenic levels and methylation levels. *limma* models adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. The solid line represents the Bonferroni threshold and the dashed line represents the FDR threshold for significance.

all 20 FDR-significant sites achieved significance at a nominal P of 0.05 for females and males. By study center, one loci was not associated with urinary As (cg20509831 located in *A1BG-AS1*; *A1BG*; *ZNF497*: nominal $P = 0.839$) and 3 loci were borderline significant (cg18616702 located in *ADAMTSL4*; *MIR4257*: nominal $P = 0.066$; cg09280971: nominal $P = 0.055$; cg07317306: nominal $P = 0.075$) in the smallest stratum representing Arizona (N = 312) while all FDR-significant loci achieved nominal significance among participants located in Oklahoma and North and South Dakota. By diabetes status, all FDR-significant loci achieved nominal significance in both strata with the exception of one CpG among participants without diabetes when using urinary creatinine concentrations to correct for urine dilution (cg14595618 located in *HK1*: nominal $P = 0.088$), and two CpGs among participants without diabetes when using specific gravity to correct for urine dilution (cg14595618 located in HK1: nominal $P = 0.283$, and cg14827056 located in *EIF2C2*: nominal $P = 0.062$).

Evaluation of previously identified differentially methylated positions

Out of 396 PubMed results containing the terms arsenic and DNA methylation, 14 reported EWAS of the association between As exposure and DNA methylation; 13 measured DNA methylation using the 450K microarray and one measured DNA methylation using the 850K microarray (Chapter 2, Table 2.4). Across these studies, 5,801 unique CpGs were associated with As exposure after adjustment for multiple comparisons (criteria used in each manuscript to determine significance are summarized in Appendix Table S7); 4,631 were included on the 850K microarray and evaluated in the current study, of which 191 achieved nominal significance in current study ($P < 0.05$). Among studies of adults, these nominally significant CpGs included cg06121226 (located in *SLC4A4*), identified as significantly associated with total blood

and urinary As levels among adults in Bangladesh (N = 400) (Argos et al., 2015), and cg05428706, cg19534475 (*ATP1B3*), and cg06466147 (*GBAP1*) identified as significantly associated with urinary or water As levels among a separate cohort of adults in Bangladesh (N = 396) (Demanelis et al., 2019). cg14718533 was found to be significantly associated with urinary As levels in a meta-analysis including both Bangladeshi cohorts (Demanelis et al., 2019) and was nominally significant in our study. In addition, cg15019001 (*HLA-DPB2*), cg22809683 (*LAMC1*), cg07466788 (*SLC16A3*), cg19504605 (ZFP41), cg22143856 (ZNF389), cg13251666, and cg13844779, identified as differentially methylated between As-exposed and control families in China (N = 102) (Guo et al., 2018), was nominally significant in our analyses. An additional 179 CpGs previously identified as associated with in utero As exposure (Cardenas et al., 2015; Gliga et al., 2018; Green et al., 2016; Kaushal et al., 2017; Kile et al., 2014; Rojas et al., 2015) were nominally significant in the current study. Demanelis et al. also reported all nominally significant loci, allowing us to check for significance of our FDR-significant CpGs. cg04940901 (*DNAH1*) and cg09280971 were associated in urinary or water As levels among the Bangladeshi cohort, and cg00500428 (*HCFC1R1*; *THOC6*) was associated with urinary As levels in the meta-analysis reported by Damanelis et al. at nominal $P < 0.05$ (Demanelis et al., 2019).

Differentially methylated regions

One DMR was identified including 20 CpGs located on chromosome 11 (chr11: 2,322,050-2,323,247) (Figure 5.3). Table 3 lists the 20 CpGs ordered by genomic coordinates. This region spans the body, first exon, and TSS200 of *C11orf21*, and the first exon, TSS200, TSS1500, and 5UTR of *TSPAN32*. Overall, the DNA methylation in this region was low, ranging from less than 1% to 12%. Fifteen of the 20 loci located within this region were positively and significantly associated with total urinary As lev-

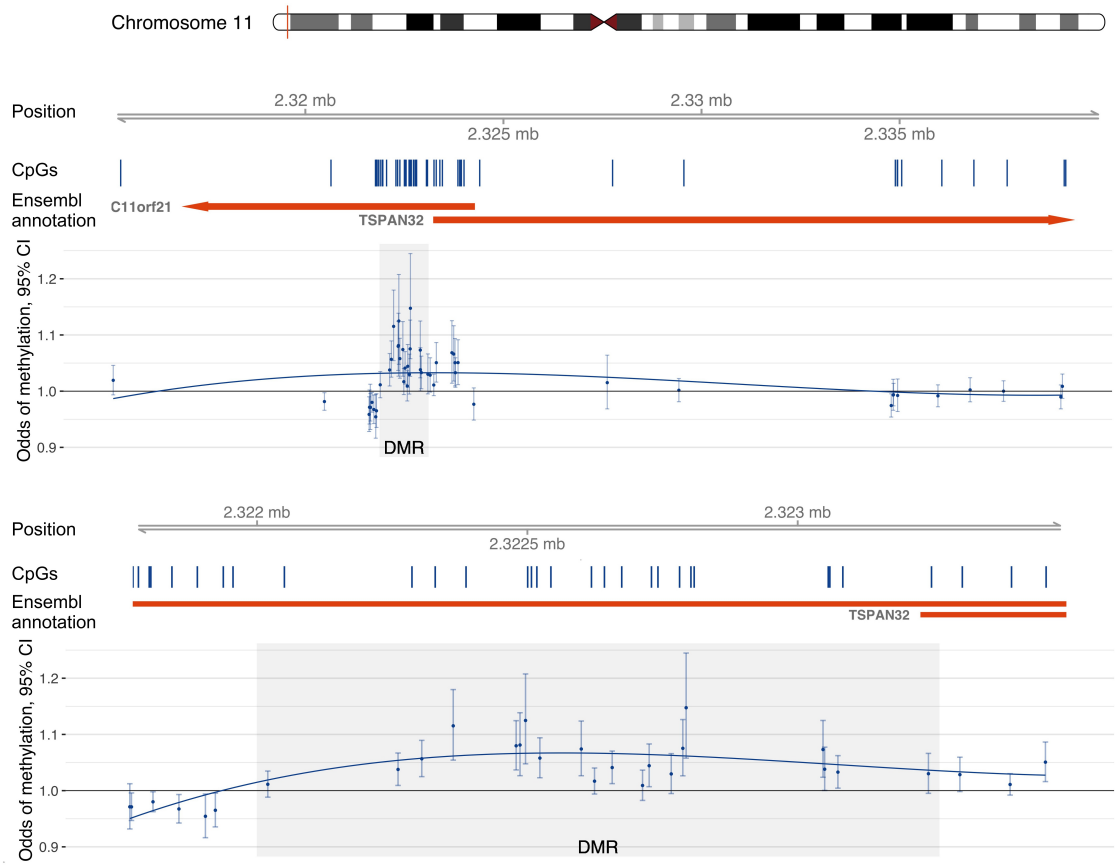


Figure 5.3: Genomic region including the differentially methylated region chr11:2,322,050-2,323,247 (shaded area). Odds of methylation and 95% CI for ln(total urinary arsenic levels) from *limma* models adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. Non-linear trend in odds of methylation related to arsenic is indicated by a solid blue line fitted using polynomial splines.

els at a nominal $P < 0.05$. With the exception of cg02537342, all CpGs were positively correlated with each other ($P < 0.05$).

Gene Ontology analysis

GO analysis identified 198 GO terms overrepresented among genes containing FDR-significant DMPs (molecular function: N = 59 GO terms; biological process: N =

Table 5.3: CpGs located in the differentially methylated region chr11:2,322,050-2,323,247.

Cpg	Position[a]	Gene	Gene region	Odds of DNAm (95% CI)[b]	*P*	Median % DNAm[c]	Mean change in % DNAm (95% CI)[d]
cg25961099	2322050	C11orf21; TSPAN32	Body; TSS1500	1.01 (0.99, 1.03)	0.019	9.02	0.08 (-0.11, 0.26)
cg01211906	2322286	C11orf21; TSPAN32	Body; TSS1500	1.04 (1.01, 1.07)	0.002	2.91	0.10 (0.02, 0.18)
cg09114153	2322329	C11orf21; TSPAN32	Body; TSS1500	1.06 (1.02, 1.09)	0.009	9.09	0.40 (0.15, 0.64)
cg01612681	2322386	C11orf21; TSPAN32	Body; TSS1500	1.12 (1.05, 1.18)	1.46×10^{-4}	1.12	0.17 (0.06, 0.29)
cg13592872	2322500	C11orf21; TSPAN32	Body; TSS1500	1.08 (1.04, 1.08)	0.502	4.2	0.28 (0.12, 0.45)
cg21027517	2322507	C11orf21; TSPAN32	Body; TSS1500	1.08 (1.03, 1.14)	0.15	3.8	0.27 (0.08, 0.46)
cg05509777	2322517	C11orf21; TSPAN32	Body; TSS1500	1.12 (1.05, 1.21)	0.096	3.75	0.37 (0.14, 0.60)
cg10782575	2322543	C11orf21; TSPAN32	Body; TSS1500	1.06 (1.02, 1.09)	0.001	10.77	0.49 (0.18, 0.81)
cg00502099	2322618	C11orf21; TSPAN32	Body; TSS1500	1.07 (1.03, 1.12)	0.092	1.18	0.09 (-0.02, 0.20)
cg03494648	2322642	C11orf21; TSPAN32	Body; TSS1500	1.02 (0.99, 1.04)	4.05×10^{-4}	1.06	0.02 (-0.02, 0.05)
cg15924868	2322674	C11orf21; TSPAN32	Body; TSS1500	1.04 (1.01, 1.07)	0.024	0.57	0.03 (0.00, 0.05)
cg02537342	2322729	C11orf21; TSPAN32	Body; TSS1500	1.01 (0.98, 1.04)	0.001	6.96	0.04 (-0.15, 0.22)
cg00041575	2322741	C11orf21; TSPAN32	Body; TSS1500	1.04 (1.01, 1.08)	2.12×10^{-4}	1.61	0.06 (-0.02, 0.15)
cg05403469	2322781	C11orf21; TSPAN32	Body; TSS1500	1.03 (0.99, 1.07)	0.048	1.34	-0.01 (-0.13, 0.12)
cg19766471	2322802	C11orf21; TSPAN32	Body; TSS1500	1.08 (1.03, 1.13)	0.005	11.75	0.72 (0.29, 1.16)
cg22210337	2322808	C11orf21; TSPAN32	Body; TSS1500	1.15 (1.06, 1.24)	0.002	11.46	1.05 (0.40, 1.70)
cg21201830	2323056	C11orf21; TSPAN32	Body, 1stExon; TSS200	1.07 (1.02, 1.12)	0.003	1.8	0.13 (0.01, 0.25)
cg15579389	2323059	C11orf21; TSPAN32	Body, 1stExon; TSS200	1.04 (1.00, 1.08)	0.003	0.89	0.04 (-0.01, 0.09)
cg09358071	2323083	C11orf21; TSPAN32	Body, 1stExon; TSS200	1.03 (1.00, 1.06)	9.33×10^{-4}	0.92	0.04 (0.01, 0.07)
cg05572370	2323247	C11orf21; TSPAN32	TSS200; 1stExon, 5'UTR	1.03 (1.00, 1.07)	0.337	2.22	0.08 (-0.02, 0.17)

a. GRCh37/hg19 assembly. b. Effect size estimate from limma models of M-values (i.e., odds of being methylated to being unmethylated for each unit change in ln(total urinary arsenic levels) adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. c. Percent methylation calculated Beta-value x 100. d. Effect size estimate from limma models of Beta-values (i.e., mean change in percent methylation for each unit change in ln(total urinary arsenic levels)) adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, >high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. DMAm, DNA methylation.

112 GO terms; cellular component: N = 27 GO terms); however, none of these terms achieved statistical significance after accounting for multiple comparisons. Table 5.4 presents the top 20 GO terms ranked by *P*-value. This list includes several sets related terms associated with the transport of cysteine. For example, GO:0015175 (neutral amino acid transmembrane transporter activity) is a parent of GO:0005294 (neutral L-amino acid secondary active transmembrane transporter activity), and GO:0005294 is a parent of GO:0015328 (cystine secondary active transmembrane transporter activity) and GO:0015327 (cystine:glutamate antiporter activity).

5.5 Discussion

This EWAS investigated the relationship between low-to-moderate levels of As exposure and loci-specific DNA methylation in a population-based prospective cohort study of American Indian adults. Twenty novel CpGs were associated with total urinary As levels ($FDR < 0.05$); methylation levels of 18 CpGs were positively associated with As levels, and methylation levels of 2 CpGs were negatively associated with As levels. In addition, one DMR located in chromosome 11 and including the genes *C11orf21* and *TSPAN32*, was identified.

The most significantly associated CpG, cg06690548, had decreased methylation levels with higher total urinary As levels. cg06690548 is located in the gene body of solute carrier family 7 member 11 (*SLC7A11*), a protein coding gene for a subunit of the amino-acid transporter cystine:glutamate antiporter system x_c^-, which exchanges cystine for glutamate within cells (Lim and Donaldson, 2011), providing cysteine for glutathione (GSH) biosynthesis (Conrad and Sato, 2012). Additionally, although not FDR significant, GO terms related to cystine:glutamate transport were identified as overrepresented among genes containing DMPs (e.g., GO:0015327, cystine:glutamate antiporter activity, and GO:0015328, cystine secondary active transmembrane trans-

Table 5.4: Top 20 gene ontology (GO) terms overrepresented among genes containing DMPs ordered by *P*-value.

Accession number	**Term**	**Ontology**[a]	**Number of genes in term**	**Number of differentially methylated genes**	***P***
GO:0015175	neutral amino acid transmembrane transporter activity	MF	33	2	7.68×10^{-4}
GO:0005294	neutral L-amino acid secondary active transmembrane transporter activity	MF	1	1	9.14×10^{-4}
GO:0015327	cystine:glutamate antiporter activity	MF	1	1	9.14×10^{-4}
GO:0015328	cystine secondary active transmembrane transporter activity	MF	1	1	9.14×10^{-4}
GO:1901680	sulfur-containing amino acid secondary active transmembrane transporter activity	MF	1	1	9.14×10^{-4}
GO:1902475	L-alpha-amino acid transmembrane transport	BP	44	2	0.001
GO:0006396	RNA processing	BP	902	5	0.002
GO:0030623	U5 snRNA binding	MF	1	1	0.002
GO:0015179	L-amino acid transmembrane transporter activity	MF	56	2	0.002
GO:0090625	mRNA cleavage involved in gene silencing by siRNA	BP	1	1	0. 002
GO:0070551	endoribonuclease activity, cleaving siRNA-paired mRNA	MF	1	1	0. 002
GO:0090624	endoribonuclease activity, cleaving miRNA-paired mRNA	MF	1	1	0. 002
GO:0015807	L-amino acid transport	BP	62	2	0. 003
GO:0036126	sperm flagellum	CC	81	2	0.003
GO:0097729	9+2 motile cilium	CC	81	2	0. 003
GO:0000386	second spliceosomal transesterification activity	MF	3	1	0. 003
GO:0000347	THO complex	CC	5	1	0. 003
GO:0000445	THO complex part of transcription export complex	CC	5	1	0. 003
GO:1905424	regulation of Wnt-mediated midbrain dopaminergic neuron differentiation	BP	2	1	0. 003
GO:1905426	positive regulation of Wnt-mediated midbrain dopaminergic neuron differentiation	BP	2	1	0. 003

a. MF = molecular function; BP = biological process; CC = cellular component.

porter activity). GSH is an endogenous antioxidant (Forman et al., 2009) and may be protective against As-induced oxidative stress. In a cross-sectional study of As-exposed adults in Bangladesh, water As concentrations were negatively associated with GSH concentrations in blood (Hall et al., 2013). Possible explanations for the observed association between As exposure and GSH concentration are that As depletes GSH through the induction of reactive oxygen species, or that As affects GSH biosynthesis through an epigenetic mechanism. Nominally significant results from GO enrichment analysis included several sets related terms associated with the transport of cysteine, further confirming the involvement of *SLC7A11A* in these pathways. A paralog of *SLC7A11*, solute carrier family 7 member 5 (*SLC7A5*) was also identified as containing a differentially methylated CpG (cg07021906). SLC7A5 is involved in the transport of amino acids including glutamine (Pochini et al., 2014), which may be deaminated to glutamate and used for GSH biosynthesis (Liu et al., 2014). SLC7A5 may also be involved in tumor promotion through the intracellular transport of leucine; leucine is an activator of mTORC1, a kinase associated with oncogenic processes (Bhutia and Ganapathy, 2016).

Additional DMPs may be biologically responsive to As exposure. Urinary As levels were positively associated with DNA methylation of cg22294740 and cg08059112, located in the 5UTR of *LINGO3*. LINGO3 is a member of the LINGO/LERN transmembrane protein family, which is commonly expressed in the nervous system, including brain tissue (Haines and Rigby, 2008). LINGO3 has also been identified as a gene hub, or key gene within a network of co-expressed genes, in metastatic melanoma tumors (Wang et al., 2018). Urinary As levels were also positively associated with DNA methylation of cg20493718 and cg21369801, located in the gene body of *CSNK1D*. The casein kinase 1 (CK1) protein family is involved in a broad range of cellular processes including control of DNA replication and repair, apoptosis, and circadian rhythm (Schittek

and Sinnberg, 2014). CK1 proteins have also been association with regulation of the tumor suppressor protein p53; specifically, CSNK1D can phosphorylate both p53 and regulatory protein MDM2 (Schittek and Sinnberg, 2014). Prenatal As exposure has been associated with differential expression (Fry et al., 2007; Rojas et al., 2015) and DNA methylation (Rojas et al., 2015) of *CSNK1D* in cord blood. In our study, As exposure was also positively associated with DNA methylation of cg14595618, located in the gene body of hexokinase 1 (*HK1*), a protein coding gene involved in glucose metabolism. Although the association between *HK1* and diabetes or diabetes-related outcomes has not been established, genetic variation in *HK1* has been associated with hemoglobin HbA1c levels, a marker of chronic glycemia (Bonnefond et al., 2009; Paré et al., 2008). Arsenic may also impact glycolysis by directly binding to enzymes involved glucose metabolism; in a study of As-binding proteins, hexokinases including HK1 were identified as As-binding proteins and As-exposed cells exhibited reduced glycolysis activity (Zhang et al., 2015).

Seven of the 20 CpGs associated with total urinary As levels at $FDR < 0.05$ were novel to the 850K (cg07317306, cg27178850, cg09280971, cg14595618, cg04940901, cg18616702, cg20509831), all of which were located in ENCODE DNase hypersensitive sites and one of which was located in a FANTOM5 enhancer (cg09280971). The ENCODE project has mapped DNase sites, associated with accessible chromatin regions (The ENCODE Project Consortium, 2012). The three significant CpGs not annotated to a RefSeq gene (cg07317306, cg27178850, cg09280971) were, however, annotated to DNase hypersensitive sites, suggesting that they may have functional roles. Two of these CpGs were also annotated to genes by the GENCODE Consortium (cg07317306 annotated to RP4-735C1.4;RP4-735C1.4 and cg27178850 annotated to RP11-467K18.2), which utilizes both manual and automated annotation processes, and has identified transcripts not included in the RefSeq database (Harrow et al., 2012).

Results were robust in sensitivity analyses stratifying by sex, study center, and diabetes status. This study includes American Indian participants recruited from three genetically, environmentally, and culturally distinct study centers in Arizona, Oklahoma, and North and South Dakota. Overall, results were consistent across study centers, providing internal validity to our analysis and suggesting that factors associated with study center do not modify the observed relationship between As exposure and DNA methylation. In addition, analyses stratified by diabetes status were consistent. Due to the effect of uncontrolled diabetes on urine osmolality (Voinescu et al., 2002), in overall analyses, we corrected for between-individual differences in urine dilution by dividing total urinary As concentrations by urinary creatinine concentrations. In analyses of participants without diabetes, total urinary As concentrations were corrected for both urinary creatinine and SG. All FDR-significant CpGs achieved significance at a nominal $P < 0.05$ among participants with diabetes. Among those without diabetes cg14595618 located in *HK1* was not significantly associated with total urinary As concentration corrected for urine dilution using urinary creatinine (nominal $P = 0.088$) or SG (nominal $P = 0.283$). We assessed possible interaction between As exposure and diabetes status on DNA methylation at cg14595618 using a linear model including the interaction term ln(total urinary As) $\times$ diabetes status and adjusted age, sex, BMI, smoking status, education, study center, eGFR, and cell type proportion estimates. Interaction between As levels and diabetes status was statistically significant ($P = 0.001$) (data not shown). Due to the role of HK1 in glucose metabolism, further research is needed to determine the relationship between As exposure, diabetes status, and *HK1* DNA methylation.

The one identified DMR (chr11:2,322,050-2,323,247) included the open reading frame *C11orf21* and spanned several functional regions of tetraspanin 32 (*TSPAN32*), including the TSS1500, TSS22, 5UTR, and 1st exon. Although no CpGs annotated to

TSPAN32 were statistically significant after adjustment for multiple comparisons in DMP analysis, 15 of the 20 CpGs in this DMR were positively associated with urinary As levels at a nominal $P < 0.05$. *C11orf21/TSPAN32* is located in a genomic region containing a cluster of imprinted genes (Smith et al., 2007) and alterations of this region have been associated with Beckwith-Wiedemann syndrome, a condition associated with abnormal growth and tumors in childhood (Koufos et al., 1989). Genetic variation in C11orf21/TSPAN32 has been associated with chronic lymphocytic leukemia in a genome-wide association study (Berndt et al., 2013).

Previous EWAS among adults have found conflicting results. Studies of adult populations have included cohorts in Bangladesh, identifying four CpGs associated with As measured in urine or blood at $P_{Bonferroni} < 0.05$ (N = 400; three CpGs overlapped for both exposure measures) (Argos et al., 2015) and 50 CpGs associated with As measured in urine or drinking water at $FDR < 0.05$ (N = 396; eight Cpgs overlapped for both exposure measures) (Demanelis et al., 2019). In addition, in a study of women in Argentina (N = 93), differential methylation was found at six loci (Ameer et al., 2017), and in a study of families in China (adults and children; N = 102), differential methylation was found at 85 loci (Guo et al., 2018). To understand the overlap between loci previously identified as differentially methylated with As exposure, we used a lookup approach of studies analyzing the association between As exposure and DNA methylation measured using the 450K or 850K microarrays. CpGs previously associated with As exposure were evaluated in the current study. Among 4,631 CpGs previously-identified as associated with As exposure after adjustment for multiple comparisons, 191 achieved nominal significance in our analyses ($P < 0.05$). This overlap between significant loci may be due to differences in population, time of exposure, tissue evaluated, and analytical methods. Of the 14 studies identified in our lookup approach, nine evaluated the association between prenatal As exposure and DNA methylation (eight measured DNA

methylation in cord blood or placental samples and one measured DNA methylation in blood mononuclear cells collected at 9 years). Only one study reported results of DNA methylation measured using the 850K microarray (Demanelis et al., 2019).

This study was limited by measuring DNA methylation in peripheral blood, and observed associations may not be present in other tissues, although As exposure is known to affect a broad range of tissues (Naujokas et al., 2013). Peripheral blood leukocytes, moreover, consist of a mixture of cell types including T cells, B cells, NK cells, monocytes, and granulocytes. Due to epigenetic control of cellular differentiation (Khavari et al., 2010), DNA methylation patterns differ between cell types, and blood cell composition may be associated with environmental exposures (Lurà et al., 2018; Stiegel et al., 2016). Therefore, EWAS using mixtures of cell types may be subject to confounding. In the current study, the Houseman regression calibration method (Houseman et al., 2012) was implemented to estimate leukocyte composition, and the proportions of six cell types were controlled for in all models. This method is generally accepted to remove confounding by cell type, although EWAS results may be influenced by variation in cell subtype proportions (Bauer et al., 2015). The study was also limited by lack of data on gene expression. Although CpGs located in biologically relevant genes were identified, it is not known if alterations in these epigenetic markers are associated with functional changes in gene expression.

The strengths of this study include the use of the 850K microarray to measure DNA methylation, the large sample size in a population with low-to-moderate levels of As exposure, and the robustness of the findings in sensitivity analyses. Previous EWAS of As exposure have predominantly measured DNA methylation using the 450K microarray, which interrogates $<$ 480,000 CpGs; the 850K microarray, however, interrogates $<$ 850,000 CpGs, including $>$ 90% of 450K loci and increased coverage of regulatory elements including ENCODE DNase hypersensitive sites and FANTOM5

enhancers (Pidsley et al., 2016). This study was also strengthened by the large sample size (N = 2,325) compared to EWAS of As exposure, which N ranged from < 50 - 400. Because power in EWAS depends on sample size, effect size, and correction for multiple testing, large sample size is particularly advantageous for 850K studies to allow for the detection of small effect sizes.

5.6 Conclusion

To our knowledge, this is the largest study to investigate the association between chronic As exposure, mostly through drinking water (Navas-Acien et al., 2009), and epigenome-wide DNA methylation in blood, and the one of the first using the 850K microarray. In a cohort of American Indian adults with low-to-moderate levels of exposure, significant associations between total urinary As levels and DNA methylation were observed at 20 novel CpGs, including loci located in genes involved in As-related mechanistic pathways and health outcomes. Further investigation is necessary to determine whether As-related DNA methylation signatures serve as biomarkers of disease development.

5.7 Bibliography

Ameer, S S; Engström, K; Hossain, M B; Concha, G; Vahter, M, and Broberg, M. Arsenic exposure from drinking water is associated with decreased gene expression and increased DNA methylation in peripheral blood. *Toxicology and Applied Pharmacology*, 321:57–66, 2017. doi: 10.1016/J.TAAP.2017.02.019.

American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*, 37 Suppl 1(Supplement 1):S81–90, 2014. doi: 10.2337/dc14-S081.

Argos, M. Arsenic exposure and epigenetic alterations: recent findings based on the Illumina 450K DNA methylation array. *Current Environmental Health Reports*, 2(2): 137–44, 2015. doi: 10.1007/s40572-015-0052-1.

Argos, M; Chen, L; Jasmine, F; Tong, L; Pierce, B L; Roy, S; Paul-Brutus, R; Gamble, M V; Harper, K N; Parvez, F; Rahman, M; Rakibuz-Zaman, M; Slavkovich, V; Baron, J A; Graziano, J H; Kibriya, M G, and Ahsan, H. Gene-specific differential DNA methylation and chronic arsenic exposure in an epigenome-wide association study of adults in Bangladesh. *Environmental Health Perspectives*, 123(1):64–71, 2015. doi: 10.1289/ehp.1307884.

Aryee, M J; Jaffe, A E; Corrada-Bravo, H; Ladd-Acosta, C; Feinberg, A P; Hansen, K D, and Irizarry, R A. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA Methylation microarrays. *Bioinformatics*, 30(10):1363–1369, 2014.

Ayotte, J D; Gronberg, J-A M, and Apodaca, L E. Trace Elements and radon in groundwater across the United States, 1992-2003. Technical report, U.S. Geological Survey Scientific Investigations Report. URL `https://water.usgs.gov/nawqa/trace/pubs/sir2011-5059/`.

Bailey, K A and Fry, R C. Arsenic-associated changes to the epigenome: what are the functional consequences? *Current Environmental Health Reports*, pages 22–34, 2014. doi: 10.1007/s40572-013-0002-8.

Bailey, K A; Smith, A H; Tokar, E J; Graziano, J H; Kim, K-W; Navasumrit, P; Ruchirawat, M; Thiantanawat, A; Suk, W A, and Fry, R C. Mechanisms underlying latent disease risk associated with early-life arsenic exposure: current research trends and scientific gaps. *Environmental Health Perspectives*, 124(2):170–175, 2016. doi: 10.1289/ehp.1409360.

Bauer, M; Linsel, G; Fink, B; Offenberg, K; Hahn, A M; Sack, U; Knaack, H; Eszlinger, M, and Herberth, G. A varying T cell subtype explains apparent tobacco smoking induced single CpG hypomethylation in whole blood. *Clinical Epigenetics*, 7(1):81, 2015. doi: 10.1186/s13148-015-0113-1.

Benjamini, Y and Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57 (1):289–300, 1995.

Berndt, S I; Skibola, C F; Joseph, V; Camp, N J; Nieters, A; Wang, Z; Cozen, W; Monnereau, A; Wang, S S, and Kelly, R S. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nature Genetics*, 45(8):868–876, 2013. doi: 10.1038/ng.2652.

Bhutia, Y D and Ganapathy, V. Glutamine transporters in mammalian cells and their functions in physiology and cancer. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1863(10):2531–2539, 2016. doi: 10.1016/J.BBAMCR.2015.12.017.

Bonnefond, A; Vaxillaire, M; Labrune, Y; Lecoeur, C; Chevre, J-C; Bouatia-Naji, N; Cauchi, S; Balkau, B; Marre, M; Tichet, J; Riveline, J-P; Hadjadj, S; Gallois, Y; Czernichow, S; Hercberg, S; Kaakinen, M; Wiesner, S; Charpentier, G; Levy-Marchal, C; Elliott, P; Jarvelin, M-R; Horber, F; Dina, C; Pedersen, O; Sladek, R; Meyre, D, and Froguel, P. Genetic variant in HK1 is associated with a proanemic state and A1C but not other glycemic control-related traits. *Diabetes*, 58(11):2687–2697, 2009. doi: 10.2337/db09-0652.

Broberg, K; Ahmed, S; Engström, K; Hossain, M B; Jurkovic Mlakar, S; Bottai, M; Grandér, M; Raqib, R, and Vahter, M. Arsenic exposure in early pregnancy alters genome-wide DNA methylation in cord blood, particularly in boys. *Journal of Developmental Origins of Health and Disease*, 5(4):288–98, 2014. doi: 10.1017/S2040174414000221.

Cardenas, A; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Wright, R O; Christiani, D C, and Kile, M L. *In utero* arsenic exposure and epigenome-wide associations in placenta, umbilical artery, and human umbilical vein endothelial cells. *Epigenetics*, 10(11):1054–1063, 2015. doi: 10.1080/15592294.2015.1105424.

Carlin, D J; Naujokas, M F; Bradham, K D; Cowden, J; Heacock, M; Henry, H F; Lee, J S; Thomas, D J; Thompson, C; Tokar, E J; Waalkes, M P; Birnbaum, L S, and Suk, W A. Arsenic and environmental health: state of the science and future research opportunities. *Environmental Health Perspectives*, 2015. doi: 10.1289/ehp.1510209.

Conrad, M and Sato, H. The oxidative stress-inducible cystine/glutamate antiporter, system x_c^-: cystine supplier and beyond. *Amino Acids*, 42(1):231–246, 2012. doi: 10.1007/s00726-011-0867-5.

Demanelis, K; Argos, M; Tong, L; Shinkle, J; Sabarinathan, M; Rakibuz-Zaman, M; Sarwar, G; Shahriar, H; Islam, T; Rahman, M; Yunus, M; Graziano, J H.; Broberg, K; Engström, K; Jasmine, F; Ahsan, H, and Pierce, B L. Association of arsenic exposure with whole blood DNA methylation: an epigenome-wide study of Bangladeshi adults. *Environmental Health Perspectives*, 127(5), 2019. doi: 10.1289/EHP3849.

Du, P; Zhang, X; Huang, C-C; Jafari, N; Kibbe, W A; Hou, L, and Lin, S M. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, 11:587, 2010. doi: 10.1186/1471-2105-11-587.

Forman, H J; Zhang, H, and Rinna, A. Glutathione: Overview of its protective roles, measurement, and biosynthesis. *Molecular Aspects of Medicine*, 30(1-2):1–12, 2009. doi: 10.1016/j.mam.2008.08.006.

Fortin, J-P; Triche, T J, and Hansen, K D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics*, 33(4): 558–560, 2017. doi: 10.1093/bioinformatics/btw691.

Fry, R C; Navasumrit, P; Valiathan, C; Svensson, J P; Hogan, B J; Luo, M; Bhattacharya, S; Kandjanapa, K; Soontararuks, S; Nookabkaew, S; Mahidol, C; Ruchirawat, M, and Samson, L D. Activation of inflammation/NF-κB signaling in infants born to arsenic-exposed mothers. *PLoS Genetics*, 3(11):e207, 2007. doi: 10.1371/journal.pgen.0030207.

Geeleher, P; Hartnett, L; Egan, L J; Golden, A; Raja Ali, R A, and Seoighe, C. Gene-set analysis is severely biased when applied to genome-wide methylation data. *Bioinformatics*, 29(15):1851–1857, 2013. doi: 10.1093/bioinformatics/btt311.

Gliga, A R; Engström, K; Kippler, M; Skröder, H; Ahmed, S; Vahter, M; Raqib, R, and Broberg, K. Prenatal arsenic exposure is associated with increased plasma IGFBP3 concentrations in 9-year-old children partly via changes in DNA methylation. *Archives of Toxicology*, 92(8):2487–2500, 2018. doi: 10.1007/s00204-018-2239-3.

Green, B B; Karagas, M R; Punshon, T; Jackson, B P; Robbins, D J; Houseman, E A, and Marsit, C J. Epigenome-wide assessment of DNA methylation in the placenta and arsenic exposure in the New Hampshire Birth Cohort Study. *Environmental Health Perspectives*, 124(8):1253–60, 2016. doi: 10.1289/ehp.1510437.

Guo, X; Chen, X; Wang, J; Liu, Z; Gaile, D; Wu, H; Yu, G; Mao, G; Yang, Z; Di, Z; Guo, X; Cao, L; Chang, P; Kang, B; Chen, J; Gao, W, and Ren, X. Multi-generational impacts of arsenic exposure on genome-wide DNA methylation and the implications for arsenic-induced skin lesions. *Environment International*, 119:250–263, 2018. doi: 10.1016/J.ENVINT.2018.06.024.

Haines, B P and Rigby, P W J. Expression of the Lingo/LERN gene family during mouse embryogenesis. *Gene Expression Patterns*, 8(2):79–86, 2008. doi: 10.1016/J.MODGEP.2007.10.003.

Hall, M N; Niedzwiecki, M; Liu, X; Harper, K N; Alam, S; Slavkovich, V; Ilievski, V; Levy, D; Siddique, A B; Parvez, F; Mey, J L; van Geen, A; Graziano, J, and Gamble, M V. Chronic arsenic exposure and blood glutathione and glutathione disulfide concentrations in Bangladeshi adults. *Environmental Health Perspectives*, 121(9): 1068–74, 2013. doi: 10.1289/ehp.1205727.

Harrow, J; Frankish, A; Gonzalez, J M; Tapanari, E; Diekhans, M; Kokocinski, F; Aken, B L; Barrell, D; Zadissa, A; Searle, S; Barnes, I; Bignell, A; Boychenko, V; Hunt, T; Kay, M; Mukherjee, G; Rajan, J; Despacio-Reyes, G; Saunders, G; Steward, C; Harte, R; Lin, M; Howald, C; Tanzer, A; Derrien, T; Chrast, J; Walters, N; Balasubramanian, S; Pei, B; Tress, M; Rodriguez, J M; Ezkurdia, I; Van Baren, J; Brent, M; Haussler, D; Kellis, M; Valencia, A; Reymond, A; Gerstein, M; Guigó, R, and Hubbard, T J. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Research*, 22(9):1760–74, 2012. doi: 10.1101/gr.135350.111.

Houseman, E A; Accomando, W P; Koestler, D C; Christensen, B C; Marsit, C J; Nelson, H H; Wiencke, J K, and Kelsey, K T. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, 13:86, 2012. doi: 10.1186/1471-2105-13-86.

IARC Working Group. Arsenic, metals, fibres, and dusts. *IARC Monographs on the Evaluation of Chronic Risks to Humans*, 100(Pt C):11–465, 2009.

Illumina. Infinium MethylationEPIC BeadChip data sheet, 2015. URL `http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/humanmethylationepic-data-sheet-1070-2015-008.pdf`.

Intarasunanont, P; Navasumrit, P; Waraprasit, S; Chaisatra, K; Suk, W A; Mahidol, C, and Ruchirawat, M. Effects of arsenic exposure on DNA methylation in cord blood samples from newborn babies and in a human lymphoblast cell line. *Environmental Health*, 11(1):31, 2012. doi: 10.1186/1476-069X-11-31.

Joint FAO/WHO Export Committee on Food Additives. Toxicological evaluation of certain food additives and contaminants. Technical report, International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland, 1989. URL `http://www.inchem.org/documents/jecfa/jecmono/v024je01.htm`.

Kaushal, A; Zhang, H; Karmaus, W J J; Everson, T M; Marsit, C J; Karagas, M R; Tsai, S-F; Wen, H-J, and Wang, S-L. Genome-wide DNA methylation at birth in relation to *in utero* arsenic exposure and the associated health in later life. *Environmental Health*, 16(1):50, 2017. doi: 10.1186/s12940-017-0262-0.

Khavari, D A; Sen, G L, and Rinn, J L. DNA methylation and epigenetic control of cellular differentiation. *Cell Cycle*, 9(19):3880–3883, 2010. doi: 10.4161/cc.9.19.13385.

Kile, M L; Baccarelli, A; Hoffman, E; Tarantini, L; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Mostofa, G; Hsueh, Y-M; Wright, R O, and Christiani, D C. Prenatal arsenic exposure and DNA methylation in maternal and umbilical cord blood leukocytes. *Environmental Health Perspectives*, 120(7):1061–1066, 2012. doi: 10.1289/ehp.1104173.

Kile, M L; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Cardenas, A; Wright, R O, and Christiani, D C. Effect of prenatal arsenic exposure on DNA methylation and leukocyte subpopulations in cord blood. *Epigenetics*, 9(5):774–82, 2014. doi: 10.4161/epi.28153.

Koufos, A; Grundy, P; Morgan, K; Aleck, K A; Hadro, T; Lampkin, B C; Kalbakji, A, and Cavenee, W K. Familial Wiedemann-Beckwith syndrome and a second Wilms tumor locus both map to 11p15.5. *American Journal of Human Genetics*, 44(5): 711–9, 1989.

Lee, E T; Welty, T K; Fabsitz, R; Cowan, L D; Le, N A; Oopik, A J; Cucchiara, A J; Savage, P J, and Howard, B V. The Strong Heart Study. A study of cardiovascular disease in American Indians: design and methods. *American Journal of Epidemiology*, 132(6):1141–55, 1990.

Leek, J T; Johnson, W E; Parker, H S; Jaffe, A E, and Storey, J D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, 28(6):882–883, 2012. doi: 10.1093/bioinformatics/bts034.

Levey, A S; Stevens, L A; Schmid, C H; Zhang, Y; Castro, A F; Feldman, H I; Kusek, J W; Eggers, P; Van Lente, F; Greene, T, and Coresh, J. A new equation to estimate glomerular filtration rate. *Annals of Internal Medicine*, 150(9):604, 2009. doi: 10.7326/0003-4819-150-9-200905050-00006.

Lim, J C and Donaldson, P J. Focus on molecules: the cystine/glutamate exchanger (System xc). *Experimental Eye Research*, 92(3):162–163, 2011. doi: 10.1016/j.exer.2010.05.007.

Liu, Y; Hyde, A S; Simpson, M A, and Barycki, J J. Emerging regulatory paradigms in glutathione metabolism. *Advances in Cancer Research*, 122:69–101, 2014. doi: 10.1016/B978-0-12-420117-0.00002-5.

Lurà, M P; Gorlanova, O; Müller, L; Proietti, E; Vienneau, D; Reppucci, D; Pavlovic, R; Dahinden, C; Röösli, MMartin; Latzin, P, and Frey, U. Response of cord blood cells to environmental, hereditary and perinatal factors: a prospective birth cohort study. *PloS one*, 13(7):e0200236, 2018. doi: 10.1371/journal.pone.0200236.

McCartney, D L; Walker, R M; Morris, S W; McIntosh, A M; Porteous, D J, and Evans, K L. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genomics Data*, 9:22–4, 2016. doi: 10.1016/j.gdata.2016.05.012.

Miller, R C; Brindle, E; Holman, D J; Shofer, J; Klein, N A; Soules, M R, and O 'connor, K A. Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. *Clinical Chemistry*, 50(5):924–932, 2004. doi: 10.1373/clinchem.2004.032292.

Moon, K A; Oberoi, S; Barchowsky, A; Chen, Y; Guallar, E; Nachman, K E; Rahman, M; Sohel, N; D'Ippoliti, D; Wade, T J; James, K A; Farzan, S F; Karagas, M R; Ahsan, H, and Navas-Acien, A. A dose-response meta-analysis of chronic arsenic exposure and incident cardiovascular disease. *International Journal of Epidemiology*, 46(6):1924–1939, 2017. doi: 10.1093/ije/dyx202.

National Research Council. Critical aspects of EPA's and IRIS assessment of inorganic arsenic, interim report. Technical report, National Academies, Washington, DC, 2013.

Naujokas, M F.; Anderson, B; Ahsan, H; Vasken Aposhian, H; Graziano, J H.; Thompson, C, and Suk, W A. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environmental Health Perspectives*, 121(3):295–302, 2013. doi: 10.1289/ehp.1205875.

Navas-Acien, A; Umans, J G; Howard, B V; Goessler, Q; Francesconi, K A; Crainiceanu, C M; Silbergeld, E K, and Guallar, E. Urine arsenic concentrations and species excretion patterns in American Indian communities over a 10-year period: The Strong Heart Study. *Environmental Health Perspectives*, 117(9):1428–1433, 2009. doi: 10.1289/ehp.0800509.

Niedzwiecki, M M; Hall, M N; Liu, X; Oka, J; Harper, K N; Slavkovich, V; Ilievski, V; Levy, D; van Geen, A; Mey, J L; Alam, S; Siddique, A B; Parvez, F; Graziano, J H, and Gamble, M V. A dose-response study of arsenic exposure and global methylation of peripheral blood mononuclear cell DNA in Bangladeshi adults. *Environmental Health Perspectives*, 121(11-12):1306–12, 2013. doi: 10.1289/ehp.1206421.

Niedzwiecki, M M; Liu, X; Hall, M N; Thomas, T; Slavkovich, V; Ilievski, V; Levy, D; Alam, S; Siddique, A B; Parvez, F; Graziano, J H, and Gamble, M V. Sex-specific associations of arsenic exposure with global DNA methylation and hydroxymethylation in leukocytes: results from two studies in Bangladesh. *Cancer Epidemiology, Biomarkers and Prevention*, 24(11):1748–57, 2015. doi: 10.1158/1055-9965.EPI-15-0432.

Paré, G; Chasman, D I; Parker, A N; Nathan, D M; Miletich, J P; Zee, R Y, and Ridker, P M. Novel association of HK1 with glycated hemoglobin in a non-diabetic population: a genome-wide evaluation of 14,618 participants in the Women's Genome Health Study. *PLoS Genetics*, 4(12):e1000312, 2008. doi: 10.1371/journal.pgen.1000312.

Peters, T J; Buckley, M J; Statham, A L; Pidsley, R; Samaras, K; Lord, R V; Clark, S J, and Molloy, P L. *De novo* identification of differentially methylated regions in the human genome. *Epigenetics and Chromatin*, 8(1):6, 2015. doi: 10.1186/1756-8935-8-6.

Phipson, B; Maksimovic, J, and Oshlack, A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics*, 32(2):btv560, sep 2015. doi: 10.1093/bioinformatics/btv560.

Pidsley, R; Zotenko, E; Peters, T J.; Lawrence, M G; Risbridger, G P; Molloy, P; Van Djik, S; Muhlhausler, B; Stirzaker, C, and Clark, S J. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology*, 17(1):208, 2016. doi: 10.1186/s13059-016-1066-1.

Pilsner, J R; Liu, X; Ahsan, H; Ilievski, V; Slavkovich, V; Levy, D; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Genomic methylation of peripheral blood leukocyte DNA: influences of arsenic and folate in Bangladeshi adults. *American Journal of Clinical Nutrition*, 86(4):1179–86, 2007.

Pilsner, J R; Liu, X; Ahsan, H; Ilievski, V; Slavkovich, V; Levy, D; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. *Environmental Health Perspectives*, 117(2):254–60, 2009. doi: 10.1289/ehp.11872.

Pochini, L; Scalise, M; Galluccio, M, and Indiveri, C. Membrane transporters for the special amino acid glutamine: structure/function relationships and relevance to human health. *Frontiers in Chemistry*, 2:61, 2014. doi: 10.3389/fchem.2014.00061.

R Core Team. R: A Language and Environment for Statistical Computing. Technical report, R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `https://www.r-project.org/`.

Ritchie, M E; Phipson, B; Wu, D; Hu, Y; Law, C W; Shi, W, and Smyth, G K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7):e47, 2015.

Robertson, K D. DNA methylation and human disease. *Nature Reviews. Genetics*, 6 (August):597–610, 2005. doi: 10.1038/nrg1655.

Roh, T; Steinmaus, C; Marshall, G; Ferreccio, C; Liaw, J, and Smith, A H. Age at exposure to arsenic in water and mortality 3040 years after exposure cessation. *American Journal of Epidemiology*, 187(11):2297–2305, 2018. doi: 10.1093/aje/kwy159.

Rojas, D; Rager, J E; Smeester, L; Bailey, K A; Drobná, Z; Rubio-Andrade, M; Stýblo, M; García-Vargas, G, and Fry, R C. Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. *Toxicological Sciences*, 143(1):97–106, 2015. doi: 10.1093/toxsci/kfu210.

Scheer, J; Findenig, S; Goessler, W; Francesconi, K A; Howard, B; Umans, J G; Pollak, J; Tellez-Plaza, M; Silbergeld, E K; Guallar, E, and Navas-Acien, A. Arsenic species and selected metals in human urine: validation of HPLC/ICPMS and ICPMS procedures for a long-term population-based epidemiological study. *Analytical Methods: Advancing Methods and Applications*, 4(2):406–413, 2012. doi: 10.1039/C2AY05638K.

Schittek, B and Sinnberg, T. Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis. *Molecular Cancer*, 13:231, 2014. doi: 10.1186/1476-4598-13-231.

Shara, N M; Wang, H; Mete, M; Al-Balha, Y R; Azalddin, N; Lee, E T; Franceschini, N; Jolly, S E; Howard, B V, and Umans, J G. Estimated GFR and incident cardiovascular disease events in American Indians: The Strong Heart Study. *American Journal of Kidney Diseases*, 60(5):795–803, 2012. doi: 10.1053/J.AJKD.2012.06.015.

Smith, A C; Choufani, S; Ferreira, J C, and Weksberg, R. Growth regulation, imprinted genes, and chromosome 11p15.5. *Pediatric Research*, 61(5 Part 2):43R–47R, 2007. doi: 10.1203/pdr.0b013e3180457660.

Smith, A H; Marshall, G; Roh, T; Ferreccio, C; Liaw, J, and Steinmaus, C. Lung, bladder, and kidney cancer mortality 40 years after arsenic exposure reduction. *Journal of the National Cancer Institute*, 110(3):241–249, 2018. doi: 10.1093/jnci/djx201.

Steinmaus, C; Ferreccio, C; Acevedo, J; Balmes, J R; Liaw, J; Troncoso, P; Dauphiné, C C; Nardone, A, and Smith, A H. High risks of lung disease associated with early-life and moderate lifetime arsenic exposure in northern Chile. *Toxicology and Applied Pharmacology*, 313:10–15, 2016. doi: 10.1016/j.taap.2016.10.006.

Stiegel, M A; Pleil, J D; Sobus, J R, and Madden, M C. Inflammatory cytokines and white blood cell counts response to environmental levels of diesel exhaust and ozone inhalation exposures. *PLOS ONE*, 11(4):e0152458, 2016. doi: 10.1371/journal.pone.0152458.

The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414):57–74, 2012. doi: 10.1038/nature11247.

U.S. Environmental Protection Agency. Government Performance and Results Act (GPRA) tool-submission year quarter 2018Q3, violation report, 2018. URL `https://obipublic.epa.gov/analytics/saw.dll?PortalPages&PortalPath=/shared/SFDW/_portal/Public&Page=Violation`.

Voinescu, G C; Shoemaker, M; Moore, H; Khanna, R, and Nolph, K D. The relationship between urine osmolality and specific gravity. *American Journal of the Medical Sciences*, 323(1):39–42, 2002.

Wang, L-X; Li, Y, and Chen, G-Z. Network-based co-expression analysis for exploring the potential diagnostic biomarkers of metastatic melanoma. *PloS one*, 13(1):e0190447, 2018. doi: 10.1371/journal.pone.0190447.

World Health Organization. Arsenic in drinking-water: background document for the development of WHO guidelines for drinking-water quality. Technical report, Geneva, Switzerland, 2011.

World Health Organization. Arsenic, 2012. URL http://www.who.int/mediacentre/factsheets/fs372/en/.

Zhang, H-N; Yang, L; Ling, J-Y; Czajkowsky, D M; Wang, J-F; Zhang, X-W; Zhou, Y-M; Ge, F; Yang, M-K; Xiong, Q; Guo, S-J; Le, H-Y; Wu, S-F; Yan, W; Liu, B; Zhu, H; Chen, Z, and Tao, S-C. Systematic identification of arsenic-binding proteins reveals that hexokinase-2 is inhibited by arsenic. *PNAS*, 112(49):15084–9, 2015. doi: 10.1073/pnas.1521316112.

5.8 Appendix

Table 5.5: Descriptive characteristics of eligible participants and participants selected for DNA methylation analysis.

	Included (N = 2,325)	**Eligible (N = 2,731)**
Age, median (IQR)	55 (49, 62)	55 (49, 62)
Sex (% male)	41.46	40.75
Smoking status		
% Current	38.41	37.79
% Former	32.17	32.95
BMI, median (IQR)	29.59 (26.22, 33.63)	29.67 (26.29, 33.69)
Education		
No high school	17.51	17.18
Some high school	23.91	23.43
Completed high school	58.58	59.39
Total urinary arsenic (μg/g creatinine)	8.56 (5.24, 14.42)	8.43 (5.15, 14.32)

Abbreviation: IQR, interquartile range.

Table 5.6: Linear models for the association between ln(total urinary arsenic) and imputed cell type proportions. Models were adjusted for age, BMI, smoking status, and sex.

Cell type	**Mean proportion**	***B* (95% CI)**	***P***
CD8+ T	0.07	0.001 (-0.004, 0.002)	0.41
CD4+ T	0.18	-0.002 (-0.006, 0.002)	0.35
NK	0.12	0.005 (0.002, 0.009)	0.004
B cells	0.08	-0.005 (-0.007, -0.002)	<0.001
Monocytes	0.05	0.002 (0.000, 0.004)	0.06
Granulocytes	0.51	0.001(-0.007, 0.008)	0.81

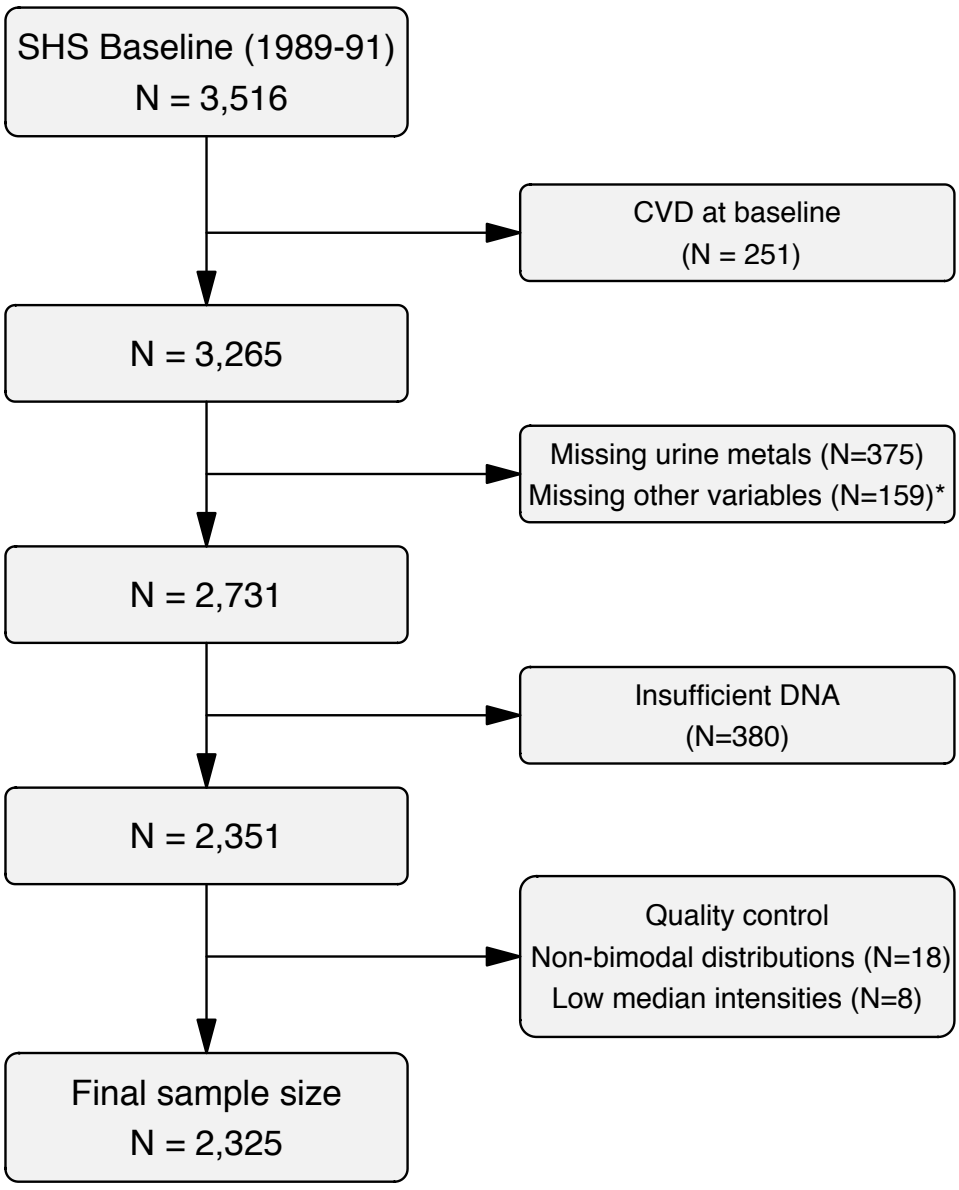


Figure 5.4: Flowchart of eligible participants and participants selected for DNA methylation analysis.

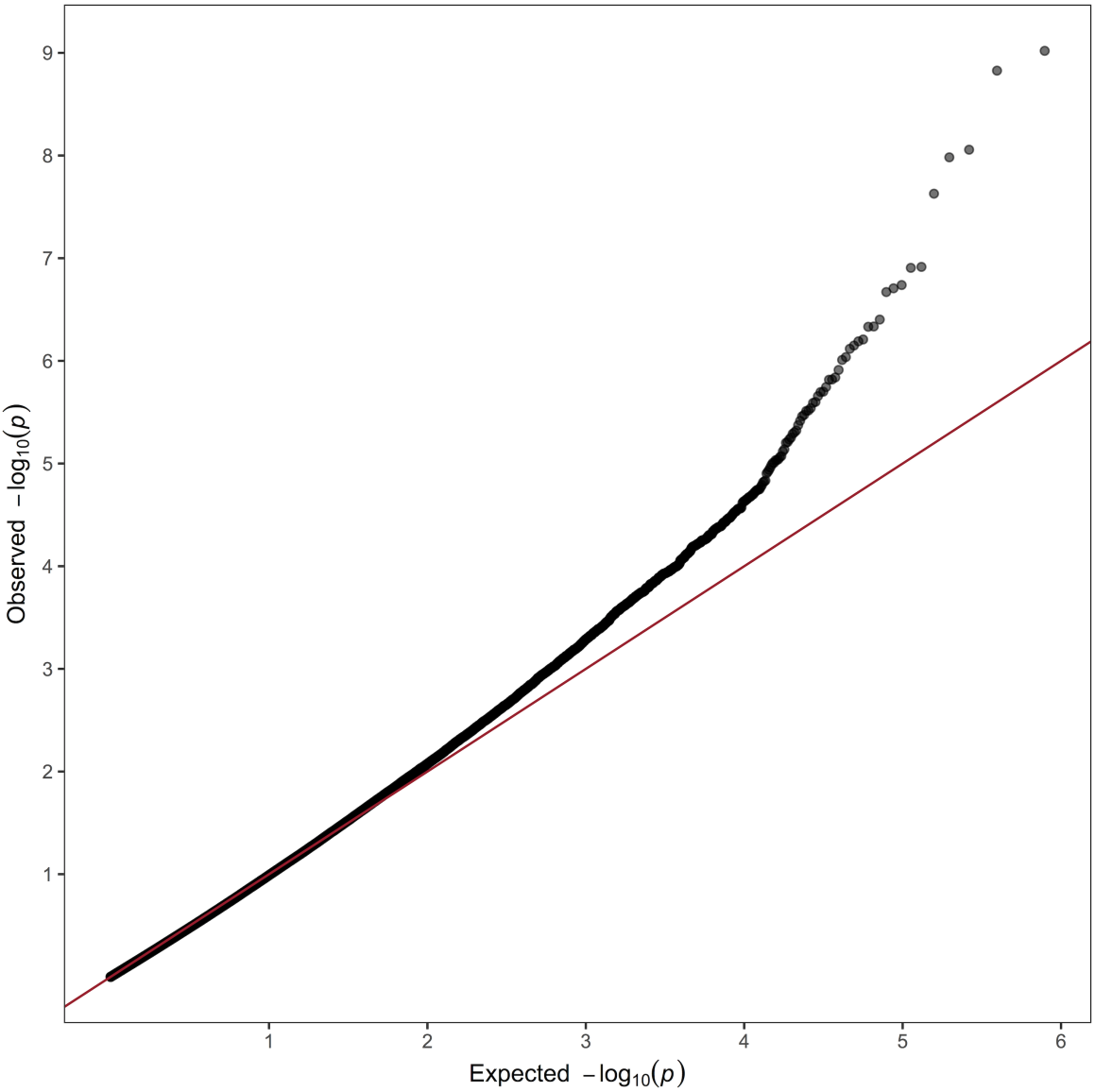


Figure 5.5: Q-Q plot from *limma* models of M-values of ln(total urinary arsenic levels) adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates.

Table 5.7: Differentially methylated positions sites associated with total urinary arsenic levels at $FDR < 0.05$, and effects sizes and P-values stratified by sex.

					Unstratified	**Female (N=1,361)**		**Male (N=964)**	
CpG	**Chr**	**Position[a]**	**Gene**	**Gene region**	***P***	**Odds of DNAm[b]**	***P***	**Odds of DNAm[b]**	***P***
cg06690548	4	139162808	SLC7A11	Body	9.56×10^{-10}	0.9	2.08×10^{-7}	0.91	2.81×10^{-4}
cg03497652	16	4751569	ANKS3	Body	1.49×10^{-9}	1.09	1.04×10^{-6}	1.09	7.22×10^{-4}
cg22294740	19	2294961	LINGO3	5'UTR	8.78×10^{-9}	1.09	2.47×10^{-6}	1.08	7.51×10^{-4}
cg20493718	17	80202979	CSNK1D	Body, 3'UTR	1.04×10^{-8}	1.14	3.24×10^{-7}	1.09	0.008
cg18616702	1	150523808	ADAMTSL4; MIR4257	5'UTR; TSS1500	2.35×10^{-8}	1.05	5.04×10^{-5}	1.06	1.66×10^{-4}
cg12116137	17	1576449	PRPF8	Body	1.22×10^{-7}	1.08	2.28×10^{-4}	1.09	7.45×10^{-4}
cg12106731	1	36023279	NCDN; KIAA0319L	TSS200; TSS1500	1.24×10^{-7}	1.04	2.62×10^{-4}	1.06	2.09×10^{-4}
cg00500428	16	3074502	HCFC1R1; THOC6	TSS1500; Body	1.83×10^{-7}	0.94	3.78×10^{-5}	0.95	0.002
cg04940901	3	52418250	DNAH1	Body	1.96×10^{-7}	1.04	0.001	1.06	5.22×10^{-5}
cg14827056	8	141550539	EIF2C2	Body	2.14×10^{-7}	1.06	4.12×10^{-4}	1.08	4.54×10^{-4}
cg21369801	17	80202961	CSNK1D	Body, 3'UTR	3.96×10^{-7}	1.08	5.91×10^{-5}	1.07	0.005
cg09280971	9	73036509	Intergenic		4.63×10^{-7}	1.02	0.002	1.04	2.40×10^{-5}
cg27178850	11	9039431	Intergenic		4.66×10^{-7}	1.03	0.001	1.05	2.45×10^{-4}
cg08059112	19	2294887	LINGO3	5'UTR	6.19×10^{-7}	1.04	3.56×10^{-5}	1.03	0.011
cg20509831	19	58866362	A1BG-AS1; A1BG; ZNF497	Body; TSS1500; 3'UTR	6.45×10^{-7}	1.04	7.03×10^{-4}	1.05	4.68×10^{-4}
cg01538969	6	30624636	DHX16	Body	7.09×10^{-7}	1.04	0.004	1.08	6.01×10^{-5}
cg14595618	10	71135446	HK1	Body	7.63×10^{-7}	1.07	3.35×10^{-5}	1.05	0.010
cg07021906	16	87866833	SLC7A5	Body	9.19×10^{-7}	1.06	0.002	1.08	3.18×10^{-4}
cg07317306	1	110314824	Intergenic		9.79×10^{-7}	1.02	0.004	1.04	2.62×10^{-4}
cg03036214	15	63640658	CA12	Body	1.23×10^{-6}	1.05	9.90×10^{-4}	1.06	0.002

a. GRCh37/hg19 assembly. b. Effect size estimate from limma models (i.e., odds of being methylated to being unmethylated for each unit change in ln(total urinary arsenic levels)) adjusted for age, sex, BMI, self-reported smoking status, education (< high school, high school graduate or GED, > high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. Abbreviations: Chr, chromosome; DNAm, DNA methylation.

Table 5.8: Differentially methylated positions sites associated with total urinary arsenic levels at $FDR < 0.05$, and effects sizes and P-values stratified by study center.

					Unstratified	**Arizona (N=312)**		**Oklahoma (N=981)**		**North & South Dakota (N=1,032)**	
CpG	**Chr**	**Position**[a]	**Gene**	**Gene region**	***P***	**Odds of DNAm**[b]	***P***	**Odds of DNAm**[b]	***P***	**Odds of DNAm**[b]	***P***
cg06690548	4	139162808	SLC7A11	Body	9.56×10^{-10}	0.9	0.026	0.87	3.42×10^{-8}	0.94	0.004
cg03497652	16	4751569	ANKS3	Body	1.49×10^{-9}	1.16	0.001	1.09	1.63×10^{-4}	1.07	8.41×10^{-7}
cg22294740	19	2294961	LINGO3	5'UTR	8.78×10^{-9}	1.18	4.73×10^{-4}	1.1	4.27×10^{-5}	1.05	0.013
cg20493718	17	80202979	CSNK1D	Body, 3'UTR	1.04×10^{-8}	1.22	0.002	1.14	1.08×10^{-4}	1.09	0.004
cg18616702	1	150523808	ADAMTSL4; MIR4257	5'UTR; TSS1500	2.35×10^{-8}	1.05	0.066	1.05	1.91×10^{-4}	1.04	0.004
cg12116137	17	1576449	PRPF8	Body	1.22×10^{-7}	1.12	0.018	1.08	0.003	1.08	0.001
cg12106731	1	36023279	NCDN; KIAA0319L	TSS200; TSS1500	1.24×10^{-7}	1.06	0.033	1.06	8.62×10^{-5}	1.04	0.005
cg00500428	16	3074502	HCFC1R1; THOC6	TSS1500; Body	1.83×10^{-7}	0.9	0.003	0.96	0.02	0.94	3.96×10^{-4}
cg04940901	3	52418250	DNAH1	Body	1.96×10^{-7}	1.08	0.026	1.04	0.01	1.05	7.17×10^{-4}
cg14827056	8	141550539	EIF2C2	Body	2.14×10^{-7}	1.1	0.027	1.06	0.003	1.05	0.004
cg21369801	17	80202961	CSNK1D	Body, 3'UTR	3.96×10^{-7}	1.16	0.002	1.07	0.004	1.06	0.01
cg09280971	9	73036509	Intergenic		4.63×10^{-7}	1.04	0.055	1.04	9.22×10^{-5}	1.02	0.018
cg27178850	11	9039431	Intergenic		4.66×10^{-7}	1.07	0.008	1.05	2.25×10^{-4}	1.02	0.028
cg08059112	19	2294887	LINGO3	5'UTR	6.19×10^{-7}	1.09	0.002	1.03	0.005	1.03	0.006
cg20509831	19	58866362	A1BG-AS1; A1BG; ZNF497	Body; TSS1500; 3'UTR	6.45×10^{-7}	1.01	0.839	1.05	0.002	1.05	6.78×10^{-5}
cg01538969	6	30624636	DHX16	Body	7.09×10^{-7}	1.13	0.002	1.05	0.006	1.04	0.016
cg14595618	10	71135446	HK1	Body	7.63×10^{-7}	1.11	0.011	1.05	0.011	1.05	0.003
cg07021906	16	87866833	SLC7A5	Body	9.19×10^{-7}	1.16	4.07×10^{-4}	1.04	0.03	1.05	0.006
cg07317306	1	110314824	Intergenic		9.79×10^{-7}	1.04	0.075	1.03	0.021	1.04	1.21×10^{-4}
cg03036214	15	63640658	CA12	Body	1.23×10^{-6}	1.09	0.02	1.05	0.01	1.04	0.008

a. GRCh37/hg19 assembly. b. Effect size estimate from limma models (i.e., odds of being methylated to being unmethylated for each unit change in ln(total urinary arsenic levels)) adjusted for age, sex, BMI, self-reported smoking status, education (<high school, high school graduate or GED, >high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. Abbreviations: Chr, chromosome; DNAm, DNA methylation.

Table 5.9: Differentially methylated positions sites associated with total urinary arsenic levels at $FDR < 0.05$, and effects sizes and P-values stratified by diabetes status.

					Unstratified	**Diabetic (N=968)**		**Non-diabetic, creatinine adjusted[c] (N=1,357)**		**Non-diabetic, SG adjusted[d] (N=1,357)**	
CpG	**Chr**	**Positiona**	**Gene**	**Gene region**	***P***	**Odds of DNAm[b]**	***P***	**Odds of DNAm[b]**	***P***	**Odds of DNAm[b]**	***P***
cg06690548	4	139162808	SLC7A11	Body	9.56×10^{-10}	0.94	0.028	0.89	3.63×10^{-9}	0.95	8.40×10^{-5}
cg03497652	16	4751569	ANKS3	Body	1.49×10^{-9}	1.11	1.24×10^{-5}	1.07	1.64×10^{-4}	0.91	1.26×10^{-4}
cg22294740	19	2294961	LINGO3	5'UTR	8.78×10^{-9}	1.1	5.78×10^{-5}	1.07	2.54×10^{-4}	1.06	2.96×10^{-4}
cg20493718	17	80202979	CSNK1D	Body, 3'UTR	1.04×10^{-8}	1.14	6.39×10^{-5}	1.11	7.53×10^{-5}	1.05	0.004
cg18616702	1	150523808	ADAMTSL4; MIR4257	5'UTR; TSS1500	2.35×10^{-8}	1.04	0.011	1.05	3.64×10^{-6}	1.1	1.03×10^{-4}
cg12116137	17	1576449	PRPF8	Body	1.22×10^{-7}	1.12	1.63×10^{-5}	1.06	0.002	1.05	1.95×10^{-6}
cg12106731	1	36023279	NCDN; KIAA0319L	TSS200; TSS1500	1.24×10^{-7}	1.05	0.001	1.05	6.24×10^{-5}	1.06	0.002
cg00500428	16	3074502	HCFC1R1; THOC6	TSS1500; Body	1.83×10^{-7}	0.95	0.007	0.94	1.78×10^{-5}	1.04	9.20×10^{-4}
cg04940901	3	52418250	DNAH1	Body	1.96×10^{-7}	1.04	0.01	1.05	4.74×10^{-5}	1.04	1.15×10^{-4}
cg14827056	8	141550539	EIF2C2	Body	2.14×10^{-7}	1.09	6.15×10^{-5}	1.05	0.003	1.03	0.062
cg21369801	17	80202961	CSNK1D	Body, 3'UTR	3.96×10^{-7}	1.09	3.32×10^{-4}	1.07	7.07×10^{-4}	1.06	0.001
cg09280971	9	73036509	Intergenic		4.63×10^{-7}	1.03	8.84×10^{-4}	1.02	0.001	1.02	0.003
cg27178850	11	9039431	Intergenic		4.66×10^{-7}	1.06	2.11×10^{-6}	1.02	0.027	1.03	0.001
cg08059112	19	2294887	LINGO3	5'UTR	6.19×10^{-7}	1.05	4.17×10^{-4}	1.03	0.002	1.02	0.008
cg20509831	19	58866362	A1BG-AS1; A1BG; ZNF497	Body; TSS1500; 3'UTR	6.45×10^{-7}	1.04	0.007	1.05	6.53×10^{-5}	1.04	4.73×10^{-4}
cg01538969	6	30624636	DHX16	Body	7.09×10^{-7}	1.07	9.09×10^{-4}	1.05	0.001	1.04	6.89×10^{-4}
cg14595618	10	71135446	HK1	Body	7.63×10^{-7}	1.11	2.72×10^{-7}	1.03	0.088	1.01	0.283
cg07021906	16	87866833	SLC7A5	Body	9.19×10^{-7}	1.09	1.41×10^{-4}	1.04	0.011	1.05	0.001
cg07317306	1	110314824	Intergenic		9.79×10^{-7}	1.03	0.004	1.03	5.75×10^{-4}	1.02	0.003
cg03036214	15	63640658	CA12	Body	1.23×10^{-6}	1.09	1.38×10^{-6}	1.03	0.036	1.03	0.028

a. GRCh37/hg19 assembly. b. Effect size estimate from limma models (i.e., odds of being methylated to being unmethylated for each unit change in ln(total urinary arsenic levels)) adjusted for age, sex, BMI, self-reported smoking status, education (<high school, high school graduate or GED, >high school), study center (Arizona, Oklahoma, North and South Dakota), estimated glomerular filtration rate, and cell type proportion estimates. c. Urinary creatinine used to adjust for urine dilution (i.e., total urinary As concentration modeled in units of μg/g creatinine). d. Specific gravity (SG) used to adjust for urine dilution. Abbreviations: Chr, chromosome; DNAm, DNA methylation.

Chapter 6

DNA methylation in cord blood as mediator of the association between prenatal arsenic exposure and gestational age

Anne K. Bozack,[1] Andres Cardenas,[2] Quazi Quamruzzaman,[3] Mahmuder Rahman,[3], Golam Mostofa,[3] David C. Christiani,[4], Molly L. Kile[5]

Affiliations: 1. Dept. of Environmental Health Sciences, Columbia University; 2. Dept. of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute; 3. Dhaka Community Hospital, Dhaka, Bangladesh; 4. Dept. of Environmental Health, Harvard T.H. Chan School of Public Health; 5. School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University

Published: https://www.ncbi.nlm.nih.gov/pubmed/30175652 Bozack et al. DNA methylation in cord blood as mediator of the association between prenatal arsenic exposure and gestational age. *Epigenetics*. (2018), doi: 10.1080/15592294.2018.1516453.

This manuscript was a collaborative effort. Anne Bozack and Andres Cardenas contributed equally to this work; analyses of the validation phase were conducted by Anne Bozack, and analyses of the discovery phase were conducted by Andreas Cardenas. The manuscript is presented in full here to provide context for this research.

6.1 Abstract

Prenatal arsenic (As) exposure is associated with adverse birth outcomes and disease risk later in life, which could be mediated through epigenetic dysregulation. We evaluated the association between As and gestational age that was mediated through DNA methylation using data from a Bangladeshi birth cohort. As exposure was measured in maternal drinking water at $\leq$ 16 weeks gestational age and maternal toenails collected $\leq$ 1 month postpartum. Cord blood DNA methylation was measured using Infinium HumanMethylation450 array (N = 44, discovery phase). Top loci identified in the discovery phase were then pyrosequenced in a second group (N = 569, validation phase). Structural equation models (SEM) evaluated the direct and indirect effects of As and DNA methylation on gestational age. In the discovery phase, As was associated with differential DNA methylation of 139 loci that were associated with gestational age ($P < 1.10\times10^{-6}$; $|\beta_{\text{regression}}| > 0.10$). Each doubling in water As concentration decreased gestational age by 2 days, which was fully mediated through the main principal component of the top-ten CpGs ($P < 0.001$). In the validation phase, there were direct and indirect effects of *miR124-3* and *MCC* DNA methylation on gestational age. In an adjusted SEM model, mediation of the association between As and gestational age by *miR124-3* was borderline significant ($P = 0.061$). This study therefore identified DNA methylation at specific loci in cord blood that mediated the effect of As exposure on gestational age. Specifically, prenatal As exposure was associated with lower methylation of *miR124-3* that mediated the exposure-response of As on gestational age. Future research should evaluate if these epigenetic changes are persistent and associated with disease risk.

Acknowledgments

We would like to thank Dr. John Geldhof from Oregon State University College of Public Health and Human Sciences for his expertise in structural equation modeling. We are also very grateful to Dr. Andres Houseman for his epigenetic expertise and insights into how to approach this analysis.

6.2 Introduction

Arsenic-contaminated drinking water is a global public health problem. Worldwide, it has been estimated that > 200 million people are chronically exposed to drinking water that contains As levels that surpass the World Health Organization and the US Environmental Protection Agency standard of 10 μg/L (Naujokas et al., 2013). Bangladesh has been particularly affected by As-contaminated drinking water, where national surveys estimate that 20 million people rely on drinking water supplies that exceed the Bangladesh national standard of 50 μg/L (Naujokas et al., 2013). This is a pressing environmental health problem because As is classified as a known Group 1 human carcinogen by the International Agency for Research on Cancer (IARC Working Group, 2009). Additionally, As readily crosses the placenta and maternal exposures are highly correlated with foetal concentrations (Concha et al., 1998). Many studies have linked maternal As exposure with adverse reproductive outcomes and adverse health effects in early childhood. For instance, prenatal As exposure has been associated with increased risk of spontaneous abortions, still birth, reduced birth weight, and both neonatal and infant mortality (Quansah et al., 2015). Prenatal As exposure has also been associated with increased susceptibility to infections during early childhood (Rahman et al., 2011; Farzan et al., 2013b) and adverse neurological and cognitive development (Quansah et al., 2015; Farzan et al., 2013b; Dangleben et al., 2013; Tyler and Allan, 2014; Cardenas et al., 2015c, 2016). Additionally, studies have also observed that As exposure early in life, particularly during gestation, increases the risk of disease and susceptibility to adverse health conditions later in life (Farzan et al., 2013a; Bailey et al., 2016).

The disruption of foetal programming events through epigenetic mechanisms has been postulated to mediate the association between environmental toxicants and health effects later in life (Saffery and Novakovic, 2014; Marsit, 2015; Perera and Herb-

stman, 2011). Epidemiological studies in adults and children report that As acts as an epigenetic toxicant and alters DNA methylation in cord or adult whole blood Argos et al. (2015); Cardenas et al. (2015a) (Argos, 2015; Cardenas et al., 2015a,b; Broberg et al., 2014; Gribble et al., 2014; Kile et al., 2012, 2014; Kaushal et al., 2017; Phookphan et al., 2017; Rojas et al., 2015). These studies provide convincing data that As exposure in the environment can alter DNA methylation in leukocytes, although there is less information linking these DNA methylation alterations to health outcomes.

In this study, our goal was to examine the association between As exposure, DNA methylation in cord blood leukocytes, and reproductive outcomes in an established prospective birth cohort recruited in Bangladesh. We utilized a two-stage approach to test the hypothesis that DNA methylation at specific loci would mediate the association between prenatal As exposure and reproductive health outcomes. This two-stage approach was economical and potentially reduced the possibility of false discoveries by coupling an agnostic epigenome-wide association study in a small discovery set with a larger candidate gene association study using pyrosequencing as a second technology.

6.3 Methods

Study population

This analysis was nested in a prospective birth cohort recruited by Dhaka Community Hospital (DCH) Trust in Bangladesh (N = 1,458). The details describing this cohort have been previously described (Kile et al., 2014). Briefly, pregnant women of $\leq$ 16 weeks of gestational age were recruited into a prospective birth cohort by DCH Trust in Bangladesh. Participants were eligible for the cohort if they had a single pregnancy, used a tube well as the main source of drinking water, and planned to live in

their current residence for the duration of the pregnancy. As part of the study protocol, women received monthly prenatal vitamins and gave birth at a local clinic or at home with DCH trained medical personnel.

This analysis was nested within the larger birth cohort. It was designed to examine the potential mediating effect of As-induced DNA methylation changes on reproductive outcomes using a more economical two-stage approach. We randomly selected 44 newborns to span a range of drinking water As concentrations (< 1-510 μg/L) for the discovery phase, which measured DNA methylation in whole cord blood with Infinium HumanMethylation450 (450K) BeadChip technology. For the validation phase, we randomly selected 569 newborns from the cohort, which measured DNA methylation in whole cord blood using pyrosequencing. Twenty-five samples from the discovery phase were also included to assess DNA methylation replication across the two technology platforms.

This study was approved by the Human Research Committees at the Harvard School of Public Health, Oregon State University, and DCH Trust.

Maternal drinking water arsenic

Arsenic was measured in tube-wells identified by participants as their main source of drinking water at the time of enrollment, as previously described (Kile et al., 2014). Briefly, water samples were collected, preserved with nitric acid to a pH < 2, and stored at room temperature prior to analysis by inductively coupled plasma-mass spectrometry (ICP-MS) using US EPA method 200.8 (Environmental Laboratory Services, North Syracuse, NY) (Creed et al., 1994). The average percent recovery for As from plasmaCal multi-element QC standard #1 solution (SCP Science) was 102 7%. The limit of detection (LOD) for As was 1 μg/L. Thirty samples were below the LOD and were assigned LOD/2.

Maternal toenail arsenic and quality control

Toenail As concentrations represent exposure during several months to year prior to collection (Kile et al., 2005). Maternal toenail clippings were collected at enrollment and ≤ 1 month after delivery. To remove contamination, toenail clippings were sonicated in 1% Triton X-100 solution (Sigma-Aldrich, Inc., St. Louis, MO) and rinsed in Milli-Q water (Millipore Corporation, Billerica, MA). Samples were digested using Trace Select Ultra Pure nitric acid (HNO3; Sigma-Aldrich, Inc.) and diluted with Milli-Q water. Total As was measured using an inductively coupled plasma mass spectrometer (Perkin-Elmer Model DRC-II 6100, Norwalk, CT). Toenail references are not available and therefore measured As concentrations were corrected for method error using blank correction and normalization based on As concentrations of batch-specific human hair references (CRM Hair; Shanghai Institute of Nuclear Research, Academia Sinica, China). Samples with a mass < 5 mg (N = 6) or relative standard deviation >25% (N = 5) were excluded from analyses. Additionally, only one remaining toenail sample was below the respective batch LOD ranging 0.004-0.85 μg As/g and subsequently excluded from the analyses.

Cord blood DNA methylation and quality control

Illumina Infinium HumanMethylation450 BeadChip: A sample of umbilical cord blood was collected after delivery into an EDTA-coated vacutainer tube (B.D. Scientific). DNA was extracted from whole blood using the Purgene DNA isolation solutions (Qiagen/Gentra Systems), following manufacturers instructions. DNA samples were analyzed for DNA methylation at the University of Minnesota Biomedical Genomic Center using the Illumina Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA), which simultaneously profiles the methylation status for $>$ 485,000 CpG loci at a single nucleotide resolution covering 99% of the RefSeq genes.

Samples were analyzed in one plate and randomly allocated to 16 chips. DNA methylation image files were normalized using the functional normalization method with two principal components to account for technical variation between samples using the minfi package of R (Aryee et al., 2014; Fortin et al., 2014). Methylation measurements at CpG loci on X and Y chromosomes were excluded from the analysis to avoid gender-specific methylation bias. Previously identified non-specific and cross-reactive probes within the array along with polymorphic CpG loci ($\geq$ 5% of the minor allele frequency) were excluded from the analysis (Chen et al., 2013). Additionally, detection *P*-values were computed for all CpGs and probes with non-significant detection (P $>$ 0.01) in greater than 10% of the samples were also excluded from the analysis. Lastly, a beta-mixture quantile intra sample normalization procedure (BMIQ) was further applied to reduce the potential bias that can arise from type-2 probes, as previously described (Teschendorff et al., 2013). The total number of autosomal loci left for analysis after quality control procedures was 383,940. Methylation values were logit-transformed to M-values to evaluate the sex adjusted linear association between CpG methylation and prenatal maternal water As exposure.

Bisulfite pyrosequencing: In the discovery phase, the top 10 CpG sites were highly correlated. Therefore, we opted to only conduct gene specific DNA methylation for the genes microRNA 124-3 (*miR124-3*), G Protein Subunit Alpha L (*GNAL*), and Mutated In Colorectal Cancers (*MCC*) (N = 569). Custom pyrosequencing was performed by EpigenDx (Hopkington, MA). Briefly, 500 ng of whole cord blood DNA was bisulfite treated using the EZ DNA Methylation kit (Zymo Research, Inc., CA). Bisulfite-treated DNA was purified following manufacturers protocols and eluted to a final volume of 46 µL. Custom PCR assays were performed using 1 µL of bisulfite-treated DNA and 0.2 µM of each primer. One primer was biotin-labeled and HPLC purified in order to purify the final PCR product using sepharose beads. PCR product

was bound to Streptavidin Sepharose HP (GE Healthcare Life Sciences) after which the immobilized PCR products were purified, washed, denatured with a 0.2 μM NaOH solution, and rewashed using the Pyrosequencing Vacuum Prep Tool (Pyrosequencing, Qiagen) following manufacturers protocol. Next, 0.5 μM of sequencing primer was annealed to the purified single stranded PCR products, and 10 μL of the PCR products were sequenced by Pyrosequencing on the PSQ96 HS System (Pyrosequencing, Qiagen) following the manufacturers instructions.

The methylation status of each CpG site was determined individually as an artificial C/T SNP using QCpG software (Pyrosequencing, Qiagen). The methylation level at each CpG site was calculated as the percentage of the methylated alleles divided by the sum of all methylated and unmethylated alleles. The mean methylation level was calculated using methylation levels of all measured CpG sites within the targeted region of each gene. Each experiment included non-CpG cytosines as internal controls to detect incomplete bisulfite conversion of the input DNA. In addition, a series of unmethylated and DNA methylation are included as controls in each PCR. The average quantity of DNA methylation and standard deviation (SD) measured at these 3 controls (e.g., low methylation (0%), medium methylation (50%), and high methylation (100%)) were estimated for each site. For *miR124-3*, the average and standard deviation (SD) of DNA methylation for the low control was 0.0 ± 0.0, medium control was 48.1 ± 22.8, and high control was 89.4 ± 10.0. For *GNAL*, the average and SD of DNA methylation for the low control was 0.07 ± 1.4, medium control was 53.2 ± 5.7, and the high control was 93.2 ± 4.7. For *MCC*, the average and SD of DNA methylation for the low control was 0.1 ± 0.7, the medium control was 64.4 ± 11.5, and the high control was 88.1 ± 7.2. Furthermore, PCR bias testing was performed by mixing unmethylated control DNA with *in vitro* methylated DNA at different ratios (0%, 5%, 10%, 25%, 50%, 75%, and 100%), followed by bisulfite modification, PCR, and pyrosequencing analysis. The

linearity of the assays were assessed using a 7 sample serial dilution ($R^2 = 0.91$ for *miRNA124-3*, 0.95 for *GNAL*, and 0.96 for *MCC*).

Statistical analysis: discovery approach

Epigenome-wide association (Step 1): A three stage filtering method was implemented to identify candidate loci that could mediate the association between prenatal As exposure and birth outcomes. The first step was to conduct an Epigenome-Wide Association Study (EWAS) in cord blood (Step 1, Figure 6.1). The EWAS in cord blood has been previously published for this sample, but we re-analyzed the data to implement the latest technical processing steps for 450K data described in the quality control section (Kile et al., 2014). After quality control, the sex adjusted linear association between maternal drinking water As and individual CpG methylation levels were evaluated using the limma function (linear models for microarray analysis) from the *minfi* package of R. A selection criteria was set *a priori* for both significance ($P < 1\times10^{-6}$) and effect size ($|\beta_{\text{regression}}| > 0.10$) to identify differentially methylated CpGs associated with As exposure *in utero* on the M-value scale. We did not adjust for cell type composition because it would not be feasible to adjust for nucleated cell type composition in the rest of the archived samples that were subsequently pyrosequenced in the validation phase due to the absence of epigenome-wide information.

Phenotype association (Step 2): All loci selected from the EWAS using the a priori criteria were then evaluated for their sex adjusted association with both birth gestational age and birth weight using multivariate linear regression models on a CpG-by-CpG basis (Step 2, Figure 6.1). Two selection criteria were used to test for association between individual CpG methylation levels and birth outcomes. Specifically, CpGs were considered to be significantly associated with birth gestational age and birth weight if: i) they reached an uncorrected level of significance of $P < 0.05$ and ii)

if multiple loci were associated only the top-ten loci ranked on lowest P-value would be selected for subsequent validation using pyrosequencing.

Mediation analysis (Step 3): Finally, CpGs identified to be differentially methylated relative to prenatal As exposure in Step 1 and found to be significantly associated with either birth gestational age or birth weight in Step 2 were evaluated for their potential to mediate the effect of exposure on these birth outcomes using structural equation models (SEMs) (Step 3, Figure 6.1). First, we checked conditions previously postulated for a variable to be considered a potential mediator (Baron and Kenny, 1986). Specifically, mediation requires a significant association between the exposure and the outcome, a significant association between the mediator and the exposure, and a significant association of the mediator to the outcome while controlling for the exposure.

We proposed a conceptual model for mediation based on the *a priori* assumption that a mediated effect through CpG DNA methylation is biologically plausible (Figure 6.2). First, we tested the independent effects of exposure on birth outcomes, exposure on CpG methylation and CpG methylation on birth outcomes while adjusting for the exposure. We then conceptualized a model in which the direct effects between prenatal As exposure, CpG methylation and birth outcomes (a, b, c) was evaluated while also testing the indirect effect of exposure on birth outcomes mediated through CpG methylation levels (a, b) while adjusting for infant sex (Figure 6.2). Two SEMs were used to evaluate the direct effect of $\log_2$-transformed maternal drinking water As on both birth gestational age and birth weight. Bias corrected standard errors and 95% bootstrap Confidence Intervals (CIs) were calculated from 10,000 replicates as the sample size available was relatively small in the discovery phase.

Histograms and scatter plots along with regression lines and locally weighted smoothing lines were plotted for bivariate association between exposure, methylation,

and birth outcomes. All pairwise Pearson correlation coefficients were evaluated among the top candidate CpGs considered for mediation analyses. Due to high correlation among all top-ten CpGs found to be associated with birth gestational age, a *post-hoc* Principal Component Analysis (PCA) was implemented to deconvolute the major source of variability into a single factor. The scores from the first principal component that accounted for the maximum amount of variability of the methylation levels of all top-ten CpGs were then evaluated as a mediator of the exposure and gestational age relationship into the conceptualized SEM model.

Statistical analysis: validation approach

Within each gene, the relationships between CpGs were assessed using Pearson correlations with the Benjamini-Hochberg false discovery rate (FDR) adjustment (Benjamini and Hochberg, 1995). SEMs were built to assess mediation between As exposure and gestational age using three steps. First, CpGs were used as indicator variables to construct a latent variable for each gene. A SNP located within *MCC* was found to be significantly associated with DNA methylation of nearby CpGs, and therefore dummy variables for the variant genotypes were included as predictors of the latent variable. Second, mediation of the association between $\log_2$-transformed maternal drinking water As concentration and birth gestational age by each latent variable (i.e., DNA methylation of each gene) was assessed. Third, a single SEM was constructed with genes that significantly mediated the association between exposure and gestational age in individual models, and adjusted for the potential confounders of sex, maternal weight gain between enrollment and delivery, maternal education ($>$ primary vs. $\leq$ primary education), and birth type (cesarean section vs. vaginal birth). Due to skew in DNA methylation variables, robust estimates of model fit were used (Brosseau-Liard et al., 2012). At each step, model indices greater than 35 were used to identify residual

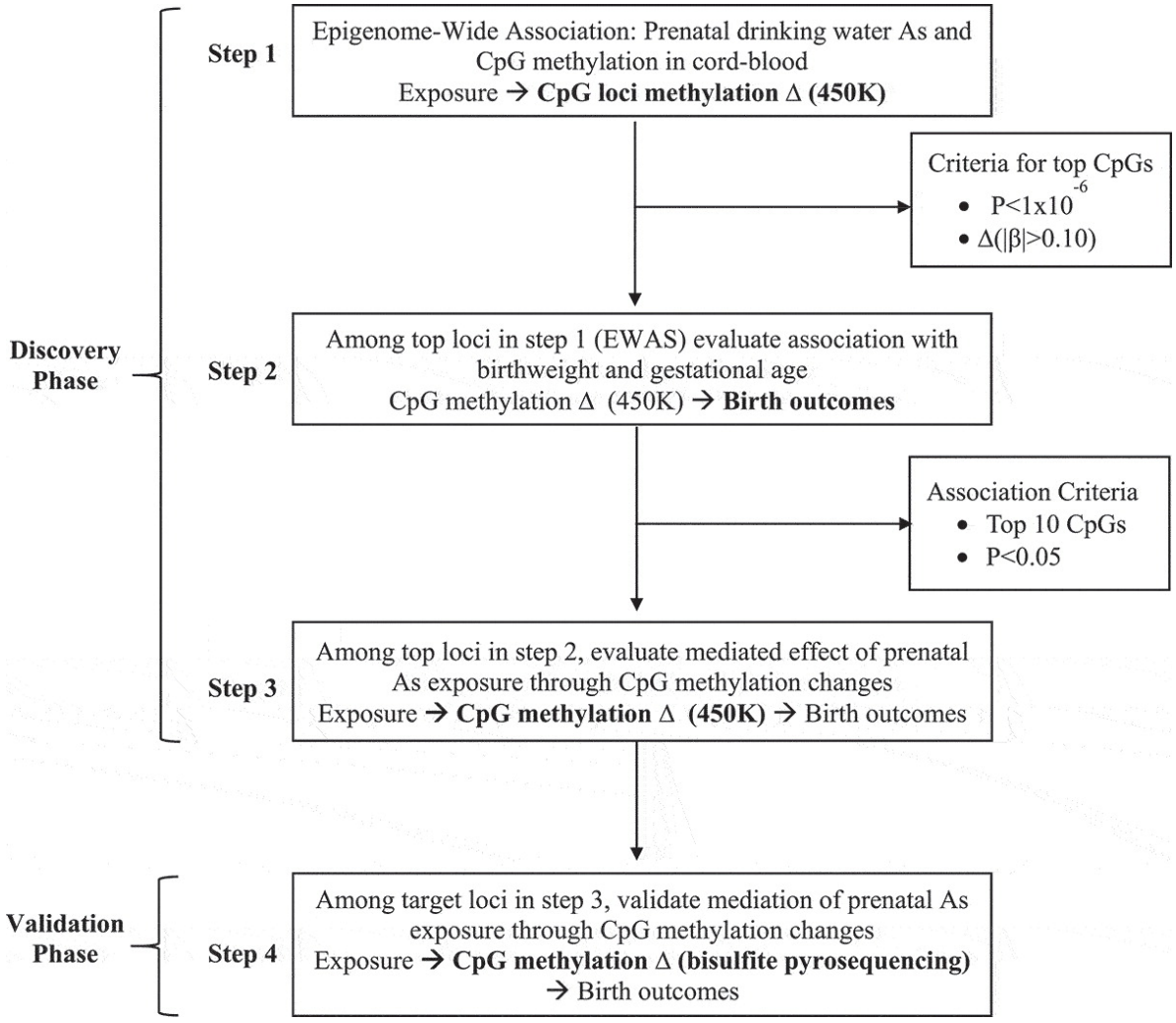


Figure 6.1: Experimental approach for the discovery and validation of DNA methylation disruption induced by prenatal As exposure and subsequent mediation of birth outcomes.

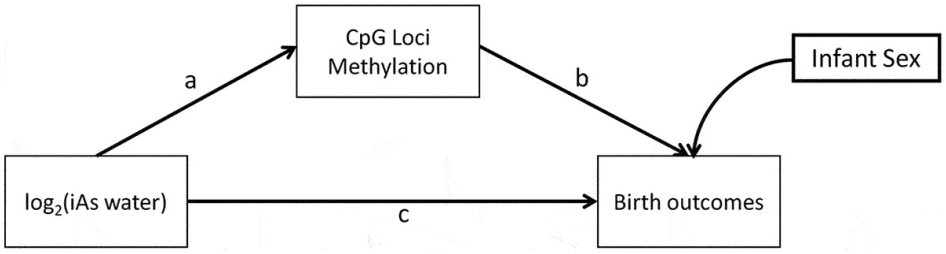


Figure 6.2: Conceptual SEM for the direct and indirect effect of exposure (maternal drinking water As $\leq$ 16 weeks of gestation) and infant birth outcomes, birth gestational age, and birth weight in the discovery phase. Abbreviation: SEM, structural equation model.

correlations that would improve model fit. All analyses were carried out using the R statistical package, version 3.4.0 or Stata, version 12.1. SEMs in R were conducted using the *lavaan* package (Oberski, 2014).

6.4 Results

Discovery phase

A total of 44 infants (29 male and 15 female) had cord blood DNA methylation measurements along with maternal drinking water As concentrations available for analysis. The mean maternal drinking water As concentration at $\leq$ 16 weeks of gestation was 63.7 μg/L (range: $<$ 1 - 510 μg/L). The mean gestational age at delivery was 37.6 weeks (range: 33 - 41 weeks) and the average birth weight was 2,923 grams (range: 2,080 - 4,050 grams). Selected sample characteristics are summarized in Table 6.1.

Table 6.1: Selected sample characteristics of the study population.

	Discovery set (N = 44)		**Validation set (N = 569)**		
Sample characteristics	**Mean ± SD**	**Range**	**Mean ± SD**	**Range**	***P*[a]**
Water As, recruitment (μg/L)	63.7 ± 116.5	<1 − 510	56.0 ± 99.8[b]	<1 − 629	0.898
Maternal toenail As, delivery (ng/μg)	7.0 ± 10.2	0.3 − 46.6	2.9 ± 3.8[c]	0.04 − 34.8	0.008
Gestational age, recruitment (weeks)	12.2 ± 2.5	6 − 16	11.4 ± 2.5	4 − 16	
Gestational age, delivery (weeks)	37.6 ± 2.1	33 − 41	37.7 ± 2.2[d]	22 − 42	0.753
Birth weight (g)	2,923 ± 372	2,080 − 4,050	2,824 ± 444[b]	1,400 − 4,600	0.141
Gender	N (%)		N (%)[b]		
Male	29 (65.9)		291 (51.2)		0.060
Female	15 (34.1)		277 (48.8)		

a. Wilcoxon rank sum test for continuous variables and Chi-squared test for categorical variables. b. N = 568. c. N = 533. d. N = 566.

Using our *a priori* selection criteria for significance of $P < 1.0\times10^6$ and effect size of $|\beta_{\text{regression}}| > 0.10$, a total of 380 loci were selected to evaluate their association with birth weight and gestational age (Figure 6.3). Among the selected 380 CpG loci identified to be differentially methylated relative to prenatal As exposure, none

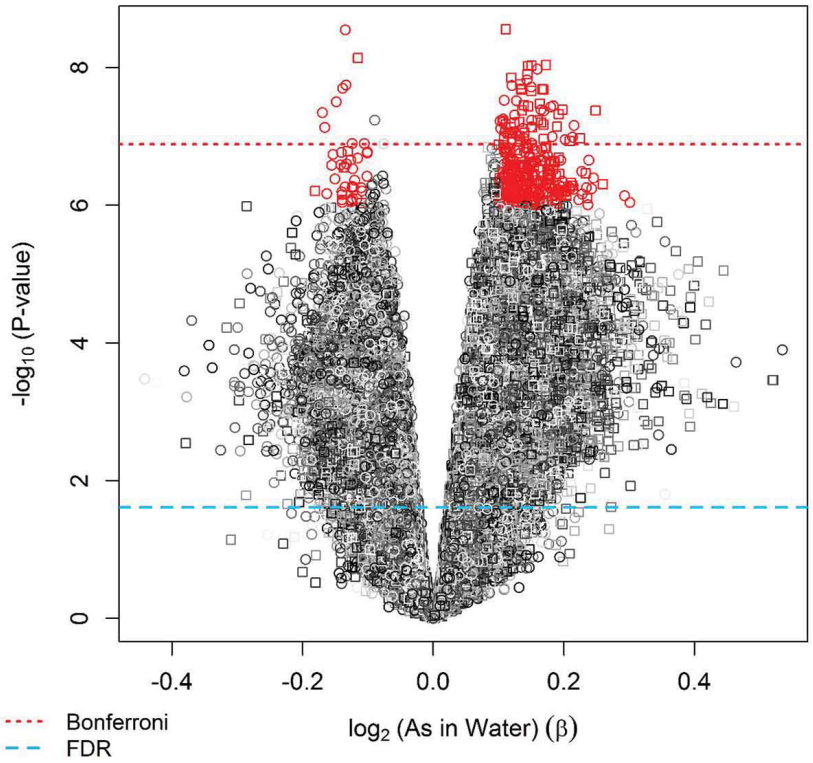


Figure 6.3: Selected differentially methylated CpG loci found in cord blood relative to prenatal As exposure using $|\beta_{\text{regression}}| < 0.10$ and nominal $P < 1.0{\times}10^{-6}$ criteria for association (Step 1 of discovery phase).

overlapped with the top 600 Houseman-probes used in differentiating white blood cell composition of whole blood samples when compared to the adult reference methylome or among the 700 CpGs used in the new cord blood reference set (Reinius et al., 2012; Bakulski et al., 2016). In sensitivity analyses, we adjusted this initial EWAS for cell type composition and, while the results were attenuated in magnitude and significance, they remained consistent (Appendix Table 6.7). We also evaluated associations with $\log_2$(postpartum maternal toenail As concentration), which are considered a very good biomarker of internal dose. The regression coefficients for the top CpGs generated from the models that used toenail As as the exposure metric were similar to models using maternal drinking water As as the measure of exposure, although significance was lower (Appendix Table 6.7).

CpG methylation and birth gestational age: Multivariate linear regression models adjusted for sex revealed that methylation levels of 139 CpGs (35.1%) from the 380 candidate loci were significantly associated with birth gestational age ($P < 0.05$) (Figure 6.4(a)). Among these loci, the top 10 CpGs ranked on lowest P-value were selected to be evaluated as mediators of the exposure and birth outcome relationship. Six of the top-ten loci were located in CpG islands and the other four in shore regions of CpG islands among unique genes and chromosomes (Table 6.2). Nine of the top-ten CpGs had higher methylation relative to prenatal As exposure and only one was observed to have lower methylation (Figure 6.5). The nine CpGs observed to be positively associated with $\log_2$-transformed As exposure (higher methylation) were inversely associated with gestational age at birth, while the single CpG loci with lower methylation was positively associated with birth gestational age (Figure 6.6).

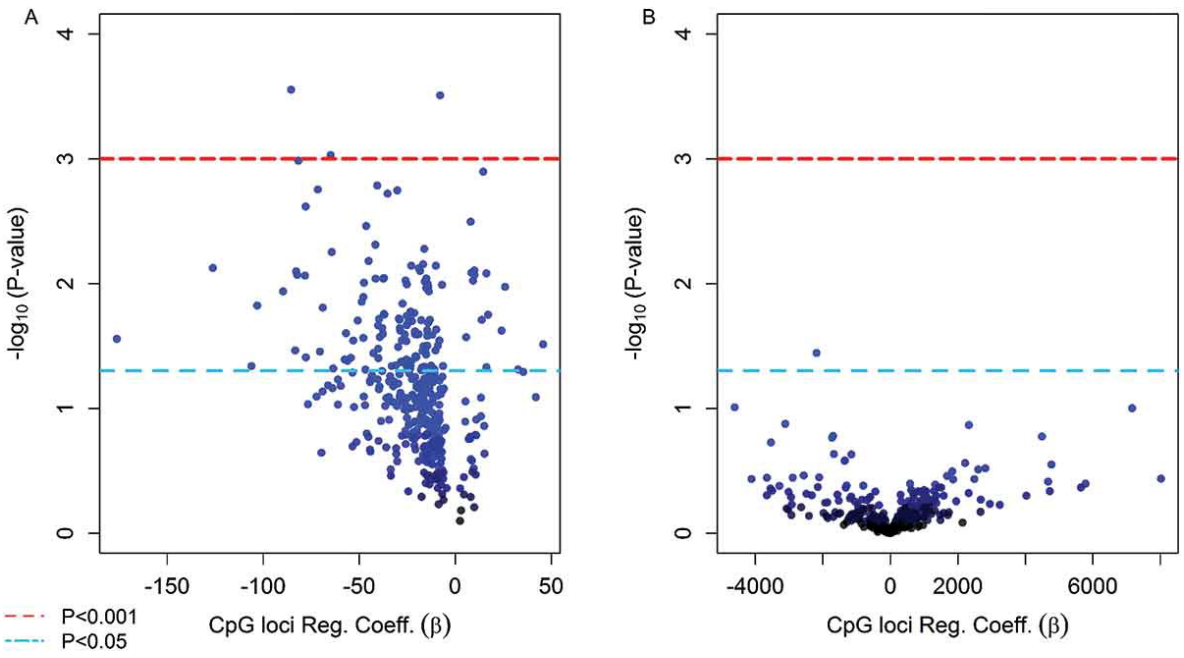


Figure 6.4: (Step 2 of discovery phase) Volcano plots for the association between the top 380 CpG loci in cord blood found to be significantly associated with As exposure *in utero* and infant health outcomes (a) gestational age and (b) birth weight.

Table 6.2: Top-ten CpG loci that were significantly associated with birth gestational age in the EWAS for the discovery phase.

CpG	*P*-value	β Coeff.	Relation to CpG Island	Chromosome	Gene	Gene Group
cg01163597	2.9×10^{-4}	−62.3	North shore	6	SLC22A23	Body
cg16081457	3.1×10^{-4}	−9.0	South shore	12		
cg06522054	3.8×10^{-4}	−68.1	Island	18	GNAL;GNAL	1stExon;Body
cg20382695	5.5×10^{-4}	−52.8	Island	10	ATRNL1	Body
cg24937280	1.1×10^{-3}	−56.7	Island	5	MCC	Body
cg01910639	1.3×10^{-3}	14.7	North shore	1	S100A6	Body
cg18115406	1.4×10^{-3}	−35.9	Island	9	LMX1B	TSS200
cg04874129	1.5×10^{-3}	−31.1	Island	16	SLC6A2	1stExon
cg20277905	1.7×10^{-3}	−39.9	Island	20	miR124-3	TSS200
cg00398764	1.7×10^{-3}	−28.0	North shore	15		

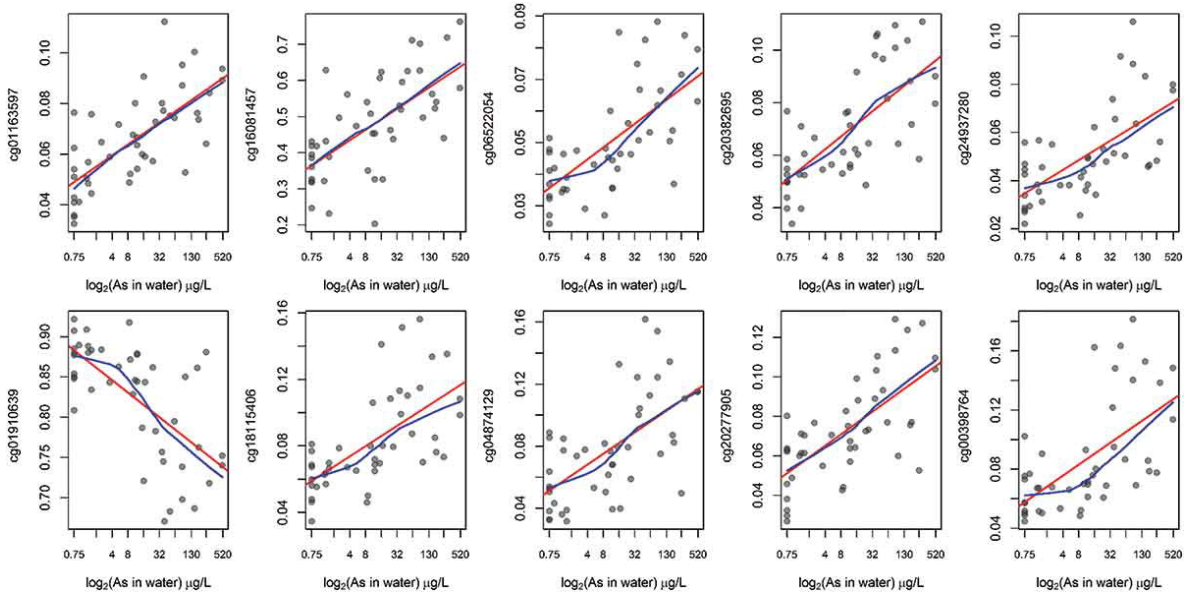


Figure 6.5: Unadjusted associations between prenatal As exposure and CpG methylation among the top-ten CpGs associated with gestational age. Red: linear regression line; blue: locally weighted scatter plot smoothing.

Nine of the top-ten CpG loci were observed to have higher methylation relative to prenatal As exposure and were positively correlated ($r_{Pearson}$ range: 0.61 - 0.90), while the single loci with lower methylation was negatively correlated with the other nine ($r_{Pearson}$ range: -0.85 - -0.68) (Figure 6.7). PCA of the top-ten loci selected demonstrated that 80% of the variance was accounted in the first and main principal

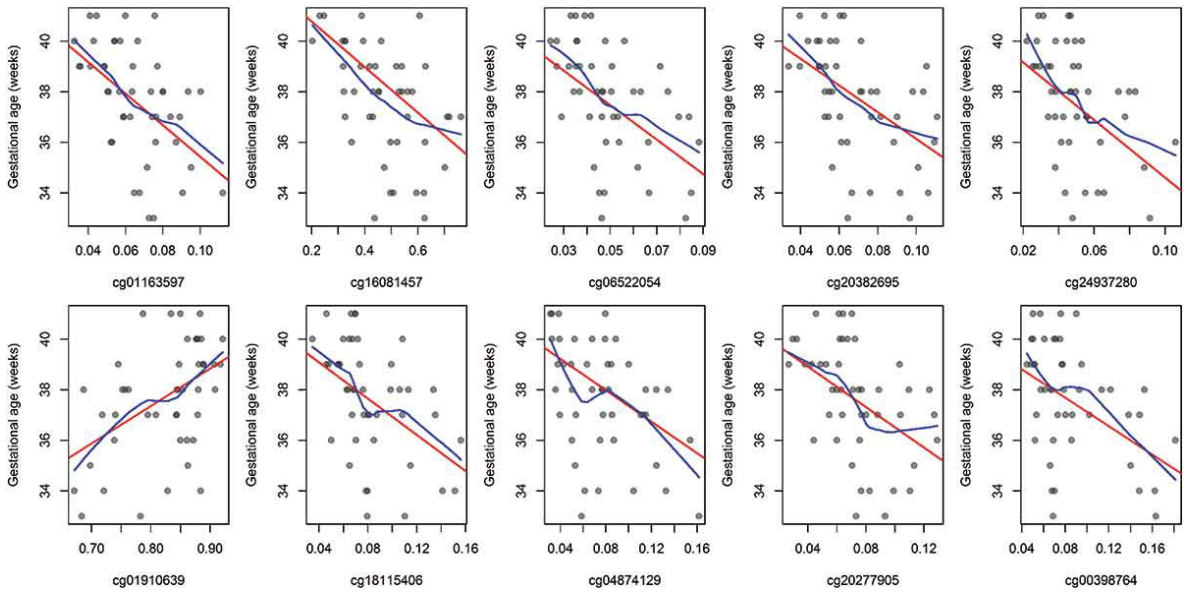


Figure 6.6: Unadjusted associations between gestational age and CpG methylation among the top-ten CpGs. Red: linear regression line; blue: locally weighted scatter plot smoothing.

component (Figure 6.8(a)) and that the cumulative variance explained by four principal components was 93% (Figure 8(b)). The relative loadings of the top-ten CpG loci on the first principal component (PC1) are presented in Appendix Table 6.8. Thus, we chose to use the first principal component due to the high level of correlation among CpGs.

Before implementing the SEM we evaluated assumptions for mediation analysis. Namely, in this subsample, $\log_2$-transformed As was significantly associated with birth gestational age (β = -0.25, 95% CI: -0.47, -0.05; $P = 0.017$) and with the scores for PC1 capturing the maximum amount of variation (80%) for the DNA methylation levels of the top-ten loci (β = 0.70, 95% CI: 0.49, 0.88; $P < 0.001$). In turn, the scores for the first principal component were significantly associated with birth gestational age while also including $\log_2$-transformed As exposure in the model (β = -0.47, 95% CI: -0.77, -0.17; $P = 0.003$), meeting the postulated conditions for mediation.

In the sex adjusted conceptual SEM $\log_2$-tranformed maternal drinking water

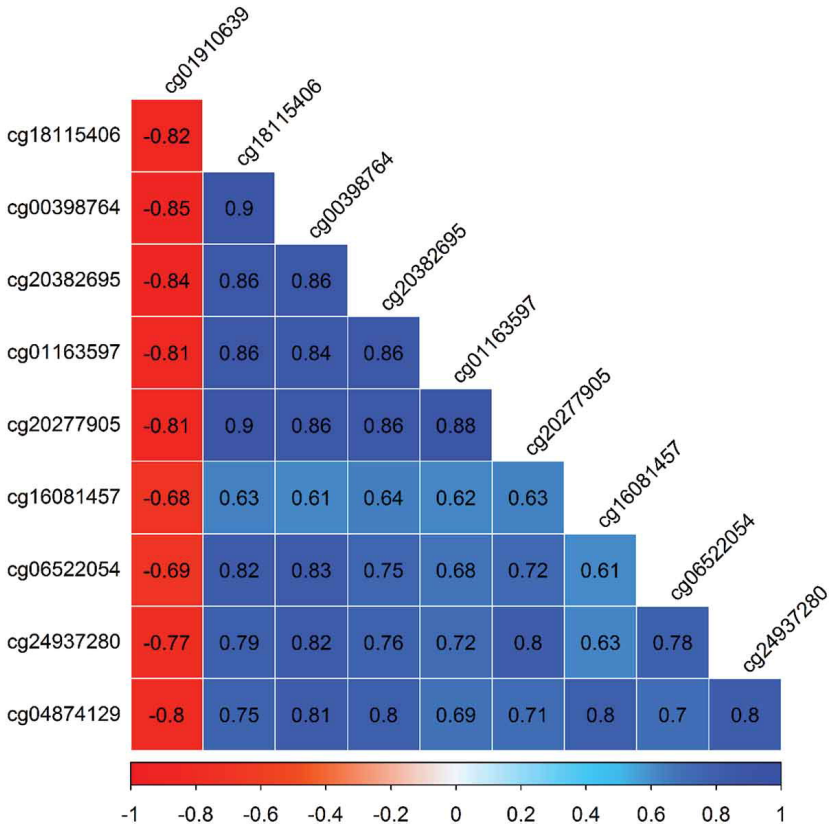


Figure 6.7: Correlation among the top-ten CpGs associated with As exposure *in utero* and gestational age at birth.

As was positively associated with PC1 scores of methylation levels for the top-ten CpGs ($\beta = 0.69$, 95% CI: 0.50, 0.87; $P < 0.001$). The principal component scores of PC1 were negatively associated with birth gestational age (β = -0.42, 95% CI: -0.59, -0.25; $P <$ 0.001) (Figure 6.9). The effect of prenatal As exposure on birth gestational age was completely mediated through PC1. Specifically, each doubling in prenatal maternal drinking water As decreased birth gestational age by 0.29 weeks or approximately two days and this was fully mediated through the PC1 scores for the methylation levels of the selected top-ten CpGs (β = -0.29, 95% CI: -0.42, -0.15; $P < 0.001$). The direct effect of maternal drinking water As on birth gestational age was non-significant after accounting for the mediation pathway and therefore not included in the final mediation model ($\beta = 0.06$, 95% CI: -0.18, 0.30; $P = 0.62$). The direct and indirect results for

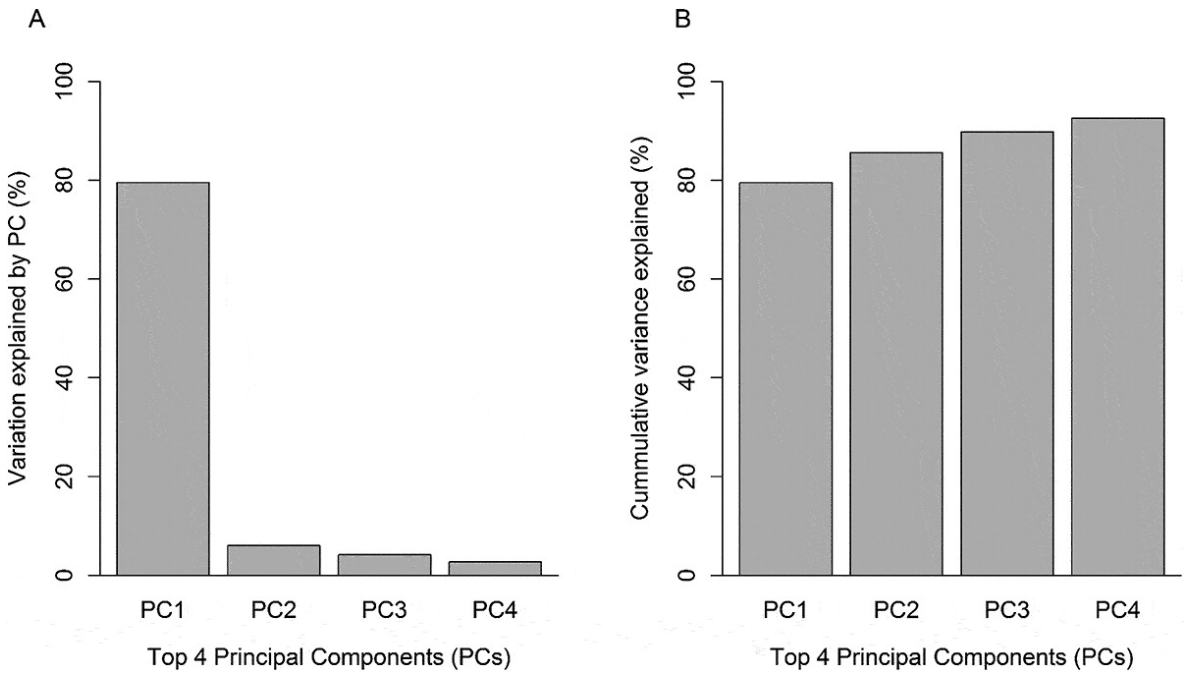


Figure 6.8: Proportion of variance explained by the first four PCs. (a) Proportion of variance explained by each PC. (b) Cumulative proportion of variance explained by all 4 PCs. Abbreviation: PC, pricipal component.

the conceptual model are summarized in Table 6.3. This final SEM conformed to all model fit indices for good fit, summarized in Table 6.4.

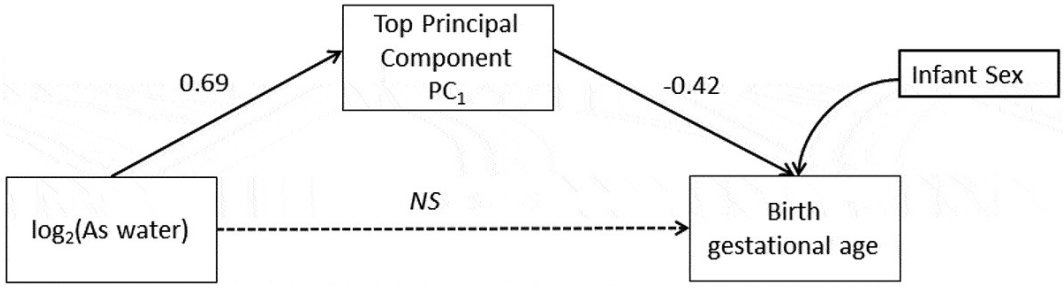


Figure 6.9: SEM conceptualized for the mediated association of the main principal component (PC1) that explained 80% of the variance for the top-10 CpG loci and birth gestational age in the discovery phase. Abbreviation: SEM, Structural equation model; PC, pricipal component.

CpG methylation and birth weight: From the 380 candidate CpGs, only one locus (cg24484905), located in an open sea region of the *DAB1* gene and observed

Table 6.3: Structural equation model for the mediated effect of As exposure on birth gestational age through variation of the top-ten CpGs captured in the main principal component (PC1) in the discovery phase.

Pathway	**Effect**	**β Coeff. (95% CIs)**	***P***
$\log_2$(Water As) → PC1	Direct	0.69 (0.50, 0.87)	<0.001
PC1 → Gestational age	Direct	-0.42 $(-0.59, -0.25)$	<0.001
Infant sex → Gestational age	Direct	-0.05 $(-1.17, 1.07)$	0.93
$\log_2$(Water As) → Gestational age	Direct	0.06 $(-0.18, 0.30)$	0.62
$\log_2$(Water As) → PC1 → Gestational age	Indirect	-0.29 $(-0.42, -0.15)$	<0.001

Table 6.4: Fit indices for the final the structural equation model that describes the indirect effect of prenatal As exposure on birth gestational age that is mediated through the main principal component (PC1) of CpG methylation levels (top-ten CpGs) in the discovery phase (N = 44).

Index	**Criterion for Good Fit**	**Model Fit**
χ^2 *P*-value	>0.05	0.81
Root Mean Square Error of Approximation (RMSEA)	<0.05	<0.001
Comparative Fit Index (CFI)	>0.95	1
Tucker-Lewisnon-normed Fit Index	>0.90	1
Standardized Root Mean Squared Residual	>0.05	0.02
Coefficient of Determination	NA	0.54

to have higher methylation relative to prenatal As exposure, was also associated with birth weight ($P = 0.035$, $FDR = 0.99$) (Figure 6.4(b)). No direct significant association was observed between maternal drinking water As and birth weight (β = 0.977, 95% CI: -27.40, 46.93; $P = 0.59$), and the direct effect of methylation levels of the *DAB1* locus (cg24484905) on birth weight was significant after controlling for As exposure (β = 2,717, 95% CI: -5,186, -248; $P = 0.032$). However, no significant mediation for the effect of prenatal As exposure on birth weight was observed through methylation levels of this single locus (β = 0.02, 95% CI: -0.01, 0.05; $P = 0.15$).

Validation stage

Bisulfite pyrosequencing was performed for target CpGs located in *miR124-3*, *GNAL*, and *MCC* on cord blood DNA from 569 infants. Among the 25 samples with Infinium 450K array and pyrosequencing data, there were strong correlations for DNA methylation measured at cg20277905 (*miR124-3*; $r_{Pearson} = 0.728$, $P < 0.001$) and cg24937280 (*MCC*; $r_{Pearson} = 0.760$; $P < 0.001$); however, correlation between platforms for cg06522054 was not significant (*MCC*; $r_{Pearson} = 0.138$; $P = 0.500$). Approximately half the infants were male (51.2% male, 48.8% female) (Table 6.1). The mean maternal drinking water As concentration at $\leq$ 16 weeks gestational age was 56.0 μg/L (range: < 1 - 629 μg/L) and the mean maternal toenail As concentration < 1 month postpartum was 2.9 ng/μg (range: 0.04 - 34.8 μg). The mean gestational age was 37.7 weeks (range: 22 - 42 weeks) and the mean birth weight was 2,824 grams (range: 1,400 - 4,600 grams). There were no significant differences in drinking water As concentration, birthweight, gestational age, or sex between participants included in the discovery and validation sets. Median postpartum maternal toenail As concentration was significantly higher among participants in the discovery set than the validation set (Wilcoxon rank sum $P = 0.008$).

In each gene, CpG loci were located within 69 - 81 base pairs and had mean methylation ranging 0.26% - 5.76% (Appendix Table 6.9). Within genes, all CpGs were significantly and positively correlated ($FDR < 0.05$) with the exception of nine CpG pairs located on *miR124-3* (Appendix Tables 6.10 - 6.12). For each gene, SEMs were used to evaluate mediation between $\log_2$-transformed maternal drinking water As and gestational age by DNA methylation (Appendix Tables 6.13 - 6.15; Appendix Figures 6.11 - 6.13). The latent variables representing DNA methylation of *miR124-3* and *MCC* were found to significantly mediate the association between prenatal As exposure and gestational age (*miR124-3* indirect effect: β = -0.02; 95% CI: -0.04, 0.00;

$P = 0.030$; *MCC* indirect effect: β = -0.03; 95% CI: -0.05, -0.01; $P = 0.004$); however, *GNAL* was not a significant mediator (*GNAL* indirect effect: β = -0.01; 95% CI: -0.02, 0.00; $P = 0.174$).

Mediation between $\log_2$-transformed maternal drinking water As and gestational age by DNA methylation of *miR124-3* and *MCC* were assessed in a single SEM. In an unadjusted model, $\log_2$-transformed maternal drinking water As concentration was significantly associated with DNA methylation of *miR124-3* and *MCC*. Furthermore, DNA methylation of *miR124-3*, but not *MCC*, was significantly associated with gestational age ($P < 0.05$) (Appendix Table 6.16; Appendix Figure 6.14). Mediation of the association between prenatal drinking water As exposure and gestational age by *miR124-3* DNA methylation achieved borderline significance (indirect effect: β = -0.02; 95% CI: -0.03, 0.00; $P = 0.051$), whereas mediation by MCC was not significant (indirect effect: β = -0.01; 95% CI: -0.03, 0.01; $P = 0.224$).

Results from a SEM adjusted for infant sex, maternal weight gain, maternal education, and birth type were consistent. There were significant direct effects of $\log_2$-transformed maternal drinking water As concentration on DNA methylation of *miR124-3* and *MCC*, and of *miR124-3* DNA methylation on gestational age ($P < 0.05$) (Figure 6.10 and Table 6.5). Mediation by DNA methylation of *miR124-3* was borderline significant (indirect effect: β = -0.02; 95% CI: -0.03, 0.00; $P = 0.061$). DNA methylation of *MCC* did not act as a mediator (indirect effect: β = -0.01; 95% CI: -0.03, 0.01; $P = 0.276$).

Sensitivity analyses were performed using maternal toenail As concentration collected at enrollment and ≤ 1 month postpartum as measures of exposure. Maternal drinking water As concentrations were significantly correlated with maternal toenail As concentrations collected at enrollment ($r_{Spearman} = 0.53,\ P < 0.001$) and postpartum ($r_{Spearman} = 0.58,\ P < 0.001$). Results from an adjusted SEM assessing mediation

Table 6.5: Adjusted structural equation model for the mediated effect of As exposure on birth gestational age through variation of DNA methylation of CpGs in *miR124-3* and *MCC* in the validation phase (N = 569).

Pathway	**Effect**	**β Coeff. (95% CIs)**	***P***
miR124-3			
$\log_2$(Water As) → *miR124-3*	Direct	0.05 (0.01, 0.10)	0.028
MCC			
$\log_2$(Water As) → *MCC*	Direct	0.06 (0.03, 0.09)	<0.001
SNP rs1057827 → *MCC* (reference: AA)			
G/G	Direct	0.53 (0.12, 0.94)	0.011
A/G	Direct	0.11 (-0.06, 0.27)	0.198
Maternal weight gain			
$\log_2$(Water As) → Weight gain	Direct	-0.15 (-0.23, -0.07)	<0.001
Birth type (Cesarean vs. vaginal) → Weight gain	Direct	0.57 (0.09, 1.09)	0.029
Gestational age			
$\log_2$(Water As) → Gestational age	Direct	-0.08 (-0.14, -0.02)	0.011
miR124-3 → Gestational age	Direct	-0.29 (-0.50, -0.09)	0.005
MCC → Gestational age	Direct	-0.17 (-0.44, 0.11)	0.242
Infant sex (Male vs. female) → Gestational age	Direct	0.08 (-0.23, 0.40)	0.600
Weight gain → Gestational age	Direct	0.23 (0.17, 0.30)	<0.001
Birth type (Cesarean section vs. vaginal) → Gestational age	Direct	0.64 (0.32, 0.95)	<0.001
Maternal education (> primary vs. ≤ primary) → Gestational age	Direct	0.47 (0.15, 0.80)	0.004
$\log_2$(Water As) → Weight gain → Gestational age	Indirect	-0.04 (-0.06, -0.02)	0.001
$\log_2$(Water As) → *miR124-3* → Gestational age	Indirect	-0.02 (-0.03, 0.00)	0.061
$\log_2$(Water As) → *MCC* → Gestational age	Indirect	-0.01 (-0.03, 0.01)	0.276

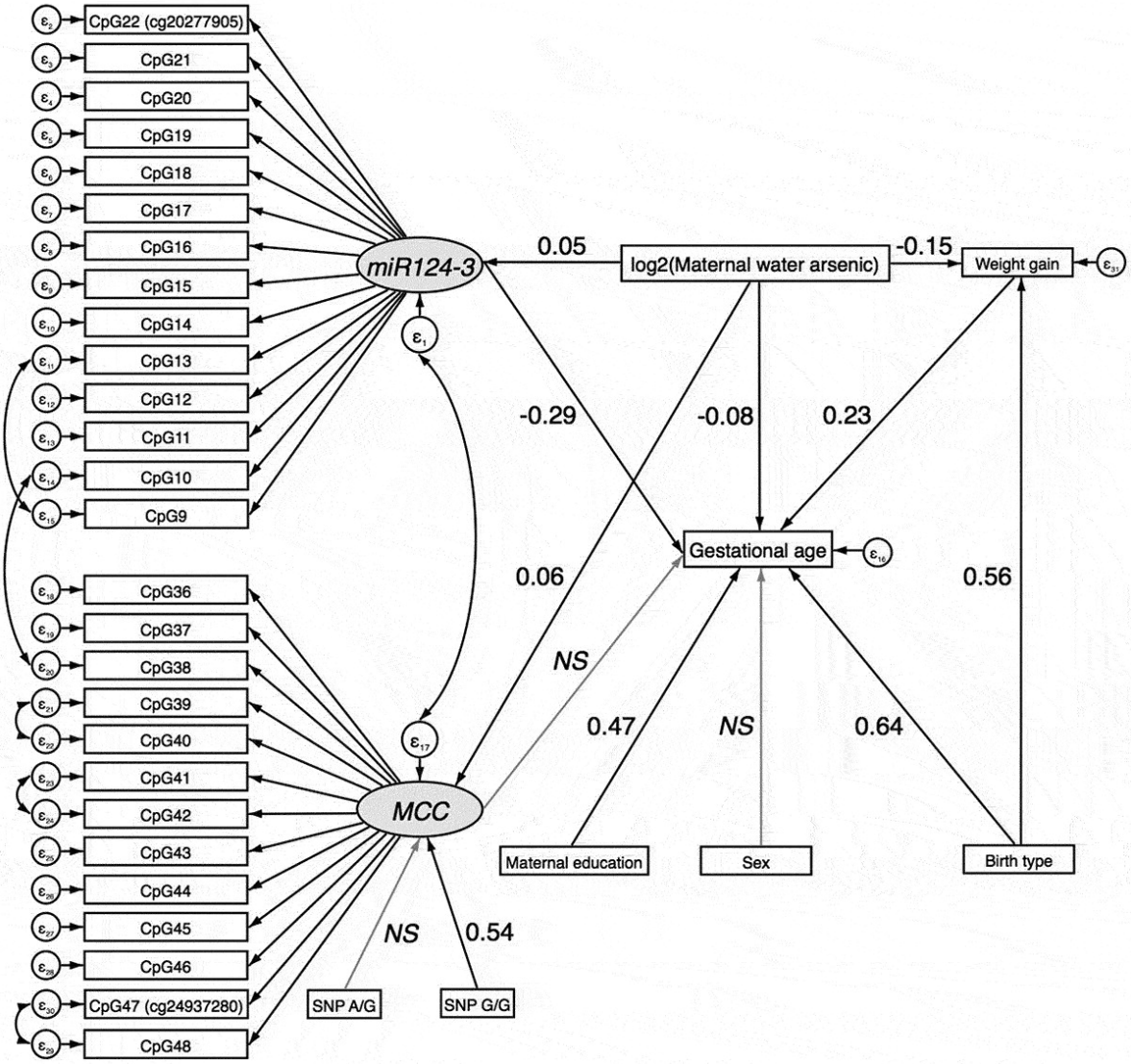


Figure 6.10: SEM for the mediated effect of As exposure on birth gestational age through DNA methylation of CpGs in *miR124-3* and *MCC* in the validation phase. Abbreviations; NS, non-significant; SEM, structural equation model.

between $\log_2$-transformed postpartum maternal toenail As and gestational age by DNA methylation of *miR124-3* and *MCC* were consistent (Appendix Figure 6.15; Appendix Table 6.17). DNA methylation of *miR124-3*, but not *MCC*, mediated with association between As exposure and gestational age (*miR124-3* indirect effect: β = -0.04; 95% CI: -0.08, -0.01; P = 0.023; *MCC* indirect effect: β = -0.02; 95% CI: -0.04, 0.01; P = 0.215). Likewise, in an adjusted SEM with $\log_2$-transformed maternal toenail

Table 6.6: Fit indices for the final the structural equation model that describe the indirect effect of prenatal arsenic exposure on birth gestational age that is mediated through variation of DNAm of CpGs in *miR124-3* and *MCC* in the validation phase (n = 569).

Index	**Criterion for Good Fit**	**Model Fit**
χ^2 *P*-value	>0.05	<0.001
Root Mean Square Error of Approximation (RMSEA)	<0.05	0.038
Comparative Fit Index (CFI)	>0.95	0.893
Tucker-Lewisnon-normed Fit Index	>0.90	0.884
Standardized Root Mean Squared Residual	>0.05	0.049

As concentration collected at enrollment, DNA methylation of *miR124-3* significantly mediated the association between prenatal As exposure and gestational age (*miR124-3* indirect effect: β = -0.05; 95% CI: -0.09, -0.01; P = 0.021; *MCC* indirect effect: β = -0.02; 95% CI: -0.05, 0.01; P = 0.210) (Appendix Figure 6.16; Appendix Table 6.18).

6.5 Discussion

We introduced an experimental approach for the discovery, evaluation, and validation of candidate CpG loci as mediators of the association between prenatal exposures and birth outcomes. Using this approach, we show that prenatal As exposure decreased birth gestational age and the association was mediated through DNA methylation levels of selected CpG loci, namely *miR124-3* and *MCC*. However, no significant mediation or direct association was observed between prenatal As exposure, DNA methylation levels at CpG loci, and birth weight.

One previous study reported an epigenome-wide association of prenatal As exposure with birth outcomes (Rojas et al., 2015). Namely, the authors observed seven unique loci significantly associated with prenatal As exposure that also correlated with birth gestational age, head circumference, or placental weight. None of the CpGs found by this group were within the top-ten differentially methylated loci found in our study

that were identified in the discovery phase. This could be potentially attributed to differences in the timing of the exposure assessment as well as the type of exposure assessment. For instance, Rojas et al. used urinary As measurements at the time of delivery, whereas we used drinking water As collected at the time of enrollment and maternal nails collected < 1 month postpartum. Furthermore, we observed that most loci were associated with gestational age and none with birth weight. This is consistent with our previously reported finding in which we show that the effect of As on birth weight is mediated through birth gestational age and to a lesser extent with maternal weight gain during pregnancy within this birth cohort (Kile et al., 2015). Another study of As exposure *in utero* using a candidate gene approach found that the expression of *AQP9*, which encodes a cell membrane channel, has the ability to mediate the effect of As on birth weight. Even though it is unknown if the *AQP9* has an epigenomic control mechanism, this observation raises the possibility that prenatal exposure can influence size at birth (Fei et al., 2013).

In our discovery phase, two CpGs (cg01163597; cg04874129) located in two genes of the solute carrier (SLC) superfamily were observed to mediate the association between prenatal As exposure and gestational age at birth (*SLC22A23*; *SLC6A2*). The *SLC6A2* is an neurotransmitter transporter across the cell membrane and has been shown to be upregulated by exposure to As in animal models (Liu et al., 2008). Furthermore, higher methylation of this gene has been associated with esophageal carcinogenesis and non-small cell lung cancer (Xu et al., 2013; Carvalho et al., 2012). The *SLC22A23* gene is a novel solute carrier protein transporter and its function has not been well characterized but abundant expression in the brain and liver has been observed (Bennett et al., 2011). It has been proposed that even though these transporters exist for endogenous substances, drugs, non-essential metals, and environmental toxins could potentially permeate. However, the physiological purpose in more than half of

these transporters remains to be characterized (He et al., 2009).

The only CpG observed to have higher methylation relative to prenatal As exposure and positively associated with gestational age was located in a north shore region of a CpG island in the body of the *S100A6* gene, involved in a Ca+2-dependent insulin release. Downregulation of this specific gene has been associated with intrauterine growth restrictions (Sitras et al., 2009). In addition, high expression levels of this protein has been observed in the human heart and, in experimental models, increased cardiac expression has been shown to be anti-hypertrophic (Tsoporis et al., 2005).

In our validation phase, candidate loci in the *miR124-3* gene were observed to have lower DNA methylation by prenatal As exposure and to mediate the association with gestational age. This specific microRNA has been correlated with tumor size and disease recurrence of non-small cell lung cancer and renal cell carcinoma (Kitano et al., 2011; Gebauer et al., 2013) and has been also shown to affect neuron growth and differentiation *in vitro* (Yu et al., 2008). The effect of prenatal exposure to As in drinking water on *miR124-3* expression has been studied in a mouse model. Tyler and Allan (2014) demonstrated that exposure to 50 μg/L As decreased *miR124-3* expression in male embryonic brain tissue (Tyler and Allan, 2014). Although *miR124-3* is predominantly expressed in the nervous system, *miR124-3* expression has been observed to regulate hematopoiesis in human cord blood cells (Liu et al., 2015). In addition, *miR124-3* may be involved in mammalian growth; fertilized mouse eggs injected with *mir124-3* microRNA resulted in increased weight at birth and in adults (Grandjean et al., 2009). There is no current evidence to suggest that *miR124-3* is imprinted. However, future work should address if *miR124-3* is a metastable epiallele potentially playing a role in the development of adult disease (Dolinoy et al., 2007).

Gestational age as an outcome is a biologically significant parameter. The clinical phenotype for early gestational age is prematurity, defined as < 37 weeks of

gestation, and preterm infants have higher rates of mortality and increased neonatal morbidity. However, recent findings also suggest a risk gradient for gestational age beyond 37 weeks compared to full term infants (Boyle et al., 2012). Reduced gestational age at birth is associated with many adverse long-term health outcomes hypothesized to be mediated by DNA methylation (Platt, 2014; Schroeder et al., 2011). Our mediation approach showed that selected CpGs had the potential to mediate the effects of prenatal exposure on gestational age. Recently few studies have started to use mediation approaches to understand the effect of environmental exposures on relevant phenotypes. For example, the effect of smoking on birth weight has been shown to be mediated by DNA methylation as well as the association between air pollution and blood pressure (Küpers et al., 2015; Bellavia et al., 2013). We further proposed that gestational age is an intermediate phenotype of disease risk later in life, potentially mediated by DNA methylation of metastable epialleles. Future prospective studies should evaluate if these epigenetic perturbations are persistent or malleable as children grow and also test if certain birth outcomes are an intermediate phenotype to for a clinical disease stage. Malleability or persistence of epigenomic modifications could yield important information on the contribution of prenatal environmental exposures to disease risk (Cardenas et al., 2017a,b).

Although no direct or mediated associations were observed between prenatal As exposure and birth weight, the CpG in the *DAB1* gene was associated with both prenatal As exposure and birth weight. *DAB1* expression has been shown to play an important role in brain ontogenesis and shown to be highly methylated in placentas of different species (Schroeder et al., 2015).

The present study has many strengths. First, the prospective measurements of the exposure, DNA methylation, and subsequent birth outcome present the possibility of testing for mediation that is chronologically possible. We also used objective

personal exposure measures during early pregnancy in which many of the fetal programming events take place. However, separate exposure measures were available in the discovery and validation phases. The discovery phase relied on a single personal water sample, which might be reasonable for the population studied. Additionally, we were able to conduct sensitivity analyses in the validation phase using postpartum maternal toenail As concentration, reflecting exposures that occurred during the gestational period, and using maternal toenail As concentration at enrollment, both of which produced similar results as personal water samples. Other studies in Bangladesh have shown that drinking water As exposures are relative constant and correlate with biomarkers of internal dose (Kile et al., 2005, 2009). Additionally, we utilized a second technology (pyrosequencing) to validate loci-specific DNA methylation discovered on the Illumina Infinium HumanMethylation450 BeadChip. This provided considerable cost-savings, as well as validation across different assays used to quantify DNA methylation in a larger set of participants. However, it should be noted that each technology has a different sensitivity for measuring DNA methylation. Women also received prenatal vitamins as part of this study and reported high compliance with taking the vitamins. Thus, micronutrient deficiencies related to vitamin B or folate are less likely to be confounders in this analysis.

There are also some important limitations to be considered. Namely, functional gene expression was not evaluated and the observed epigenetic disruption might not lead to physiological changes in expression. Although the unadjusted EWAS was used for the identification of differentially methylated loci, the potential of confounding by shifts in white blood cell composition was minimized by ensuring that the probes selected did not differentiate cell types when using the Houseman method and available reference panels. Therefore, results should be interpreted in light of this and could indeed reflect cell type distribution. In addition, in the discovery phase, our relatively

small sample size is an important limitation as it does not allow us to adjust for other potential confounders. However, use of a larger set of the birth cohort in the validation phase allowed for adjustment of multiple confounders. It should also be noted that maternal toenail As concentrations, which can serve as a more accurate measure of exposure than drinking water As concentrations, were not available when the discovery phase was completed. However, we performed sensitivity analyses in the validation phase using As concentrations from maternal toenail samples collected at enrollment and postpartum. Overall, the results of these analyses were consistent. The significance of mediation of the association between prenatal As exposure and gestational age by DNA methylation of *miR124-3* increased with use of toenail As as the measure of exposure.

6.6 Conclusion

In summary, in a two-stage experimental approach using discovery and validation phases we show that prenatal As exposure is inversely associated with birth GA and the association mediated by DNAm of *miR124-3* and *MCC*. However, no direct or mediated association was observed for birth weight. Our results support the hypothesis that arsenic exposure in utero can disrupt fetal programming leading to phenotypic consequences that may play a role in the developmental origins of health and disease. Furthermore, this experimental framework for the discovery and validation of candidate CpG loci as mediators of exposures and health outcomes could be extended to other exposures and health outcomes.

6.7 Bibliography

Argos, M. Arsenic exposure and epigenetic alterations: recent findings based on the Illumina 450K DNA methylation array. *Current Environmental Health Reports*, 2(2): 137–44, 2015. doi: 10.1007/s40572-015-0052-1.

Argos, M; Chen, L; Jasmine, F; Tong, L; Pierce, B L; Roy, S; Paul-Brutus, R; Gamble, M V; Harper, K N; Parvez, F; Rahman, M; Rakibuz-Zaman, M; Slavkovich, V; Baron, J A; Graziano, J H; Kibriya, M G, and Ahsan, H. Gene-specific differential DNA methylation and chronic arsenic exposure in an epigenome-wide association study of adults in Bangladesh. *Environmental Health Perspectives*, 123(1):64–71, 2015. doi: 10.1289/ehp.1307884.

Aryee, M J; Jaffe, A E; Corrada-Bravo, H; Ladd-Acosta, C; Feinberg, A P; Hansen, K D, and Irizarry, R A. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA Methylation microarrays. *Bioinformatics*, 30(10):1363–1369, 2014.

Bailey, K A; Smith, A H; Tokar, E J; Graziano, J H; Kim, K-W; Navasumrit, P; Ruchirawat, M; Thiantanawat, A; Suk, W A, and Fry, R C. Mechanisms underlying latent disease risk associated with early-life arsenic exposure: current research trends and scientific gaps. *Environmental Health Perspectives*, 124(2):170–175, 2016. doi: 10.1289/ehp.1409360.

Bakulski, K M; Feinberg, J I; Yang, J; Brown, S; Andrews, S V; McKenney, S; Witter, F; Walston, Feinberg, J, and AP, Fallin M D. DNA methylation of cord blood cell types: Applications for mixed cell birth studies. *Epigenetics*, 11(5):354–62, 2016.

Baron, R M and Kenny, D A. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182, 1986.

Bellavia, A; Urch, B; Speck, M; Brook, R D; Scott, J A; Albetti, B; Behbod, B; North, M; Valeri, L; Bertazzi, P A; Silverman, F; Gold, D, and Baccarelli, A A. DNA hypomethylation, ambient particulate matter, and increased blood pressure: findings from controlled Human exposure experiments. *Journal of the American Heart Association*, 2(3):e000212, 2013. doi: 10.1161/JAHA.113.000212.

Benjamini, Y and Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57 (1):289–300, 1995.

Bennett, K M; Liu, J; Hoelting, C, and Stoll, J. Expression and analysis of two novel rat organic cation transporter homologs, SLC22A17 and SLC22A23. *Molecular and Cellular Biochemistry*, 352(1-2):143–154, 2011. doi: 10.1007/s11010-011-0748-y.

Boyle, E M; Poulsen, G; Field, D J; Kurinczuk, J J; Wolke, D; Alfirevic, Z, and Quigley, M A. Effects of gestational age at birth on health outcomes at 3 and 5 years of age: population based cohort study. *BMJ*, 344:e896, 2012. doi: 10.1136/BMJ.E896.

Broberg, K; Ahmed, S; Engström, K; Hossain, M B; Jurkovic Mlakar, S; Bottai, M; Grandér, M; Raqib, R, and Vahter, M. Arsenic exposure in early pregnancy alters genome-wide DNA methylation in cord blood, particularly in boys. *Journal of Developmental Origins of Health and Disease*, 5(4):288–98, 2014. doi: 10.1017/S2040174414000221.

Brosseau-Liard, Patricia E; Savalei, V, and Li, L. An investigation of the sample performance of two nonnormality corrections for RMSEA. *Multivariate Behavioral Research*, 47(6):904–930, 2012. doi: 10.1080/00273171.2012.715252.

Cardenas, A; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Wright, R O; Christiani, D C, and Kile, M L. *In utero* arsenic exposure and epigenome-wide associations in placenta, umbilical artery, and human umbilical vein endothelial cells. *Epigenetics*, 10(11):1054–1063, 2015a. doi: 10.1080/15592294.2015.1105424.

Cardenas, A; Koestler, D C; Houseman, E A; Jackson, B P; Kile, M L; Karagas, M R, and Marsit, C J. Differential DNA methylation in umbilical cord blood of infants exposed to mercury and arsenic *in utero*. *Epigenetics*, 10(6):508–515, 2015b. doi: 10.1080/15592294.2015.1046026.

Cardenas, A; Smit, E; Houseman, E A; Kerkvliet, N I; Bethel, J W, and Kile, M L. Arsenic Exposure and Prevalence of the Varicella Zoster Virus in the United States: NHANES (20032004 and 20092010). *Environmental Health Perspectives*, 123(6):590–596, 2015c. doi: 10.1289/ehp.1408731.

Cardenas, A; Smit, E; Bethel, J W; Houseman, E A, and Kile, M L. Arsenic exposure and the seroprevalence of total hepatitis A antibodies in the US population: NHANES, 20032012. *Epidemiology and Infection*, 144(8):1641–1651, 2016. doi: 10.1017/S0950268815003088.

Cardenas, A; Rifas-Shiman, S L; Agha, G; Hivert, M-F; Litonjua, A A; DeMeo, D L; Lin, X; Amarasiriwardena, C J; Oken, E; Gillman, M W, and Baccarelli, A A. Persistent DNA methylation changes associated with prenatal mercury exposure and cognitive performance during childhood. *Scientific Reports*, 7(1):288, 2017a. doi: 10.1038/s41598-017-00384-5.

Cardenas, A; Rifas-Shiman, S L; Godderis, L; Duca, R-C; Navas-Acien, A; Litonjua, A A.; DeMeo, D L; Brennan, K J; Amarasiriwardena, C J; Hivert, M-F; Gillman, M W; Oken, E, and Baccarelli, A A. Prenatal exposure to mercury: associations with global DNA methylation and hydroxymethylation in cord blood and in childhood. *Environmental Health Perspectives*, 125(8):087022, 2017b. doi: 10.1289/EHP1467.

Carvalho, R; Haberle, V; Hou, J; van Gent, T; Thongjuea, S; van IJcken, W; Kockx, C; Brouwer, R; Rijkers, E; Sieuwerts, A; Foekens, J; van Vroonhoven, M; Aerts, J; Grosveld, F; Lenhard, B, and Philipsen, S. Genome-wide DNA methylation profiling of non-small cell lung carcinomas. *Epigenetics and Chromatin*, 5(1):9, 2012. doi: 10.1186/1756-8935-5-9.

Chen, Y-A; Lemire, M; Choufani, S; Butcher, D T; Grafodatskaya, D; Zanke, B W; Gallinger, S; Hudson, T J, and Weksberg, R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*, 8(2):203–9, 2013. doi: 10.4161/epi.23470.

Concha, G; Vogler, G; Lezcano, D; Nermell, B, and Vahter, M. Exposure to inorganic arsenic metabolites during early human development. *Toxicological Sciences*, 44(2): 185–190, 1998.

Creed, J; Brockhoff, C, and Martin, T. Method 200.8, Revision 5.4: Determination of trace elements in waters and wastes by inductively coupled plasma mass spectrometry. Technical report, Environmental Monitoring Systems Laboratory, Office of Research and Development, US Environmental Protection Agency, Cincinnati, OH, 1994. URL `www.epa.gov`.

Dangleben, Nygerma L; Skibola, Christine F, and Smith, Martyn T. Arsenic immunotoxicity: a review. *Environmental Health*, 12(1):73, 2013. doi: 10.1186/1476-069X-12-73.

Dolinoy, D C; Das, R; Weidman, J R, and Jirtle, R L. Metastable epialleles, imprinting, and the fetal origins of adult diseases. *Pediatric Research*, 61(5 Part 2):30R–37R, 2007. doi: 10.1203/pdr.0b013e31804575f7.

Farzan, S F; Karagas, M R, and Chen, Y. *In utero* and early life arsenic exposure in relation to long-term health and disease. *Toxicology and Applied Pharmacology*, 272 (2):384–390, 2013a. doi: 10.1016/j.taap.2013.06.030.

Farzan, S F.; Korrick, S; Li, Z; Enelow, R; Gandolfi, A J; Madan, J; Nadeau, K, and Karagas, M R. *In utero* arsenic exposure and infant infection in a United States cohort: a prospective study. *Environmental Research*, 126:24–30, 2013b. doi: 10.1016/j.envres.2013.05.001.

Fei, D L; Koestler, D C; Li, Z; Giambelli, C; Sanchez-Mejias, A; Gosse, J A; Marsit, C J; Karagas, M R, and Robbins, D J. Association between *in utero* arsenic exposure, placental gene expression, and infant birth weight: a US birth cohort study. *Environmental Health*, 12(1):58, 2013. doi: 10.1186/1476-069X-12-58.

Fortin, J-P; Labbe, A; Lemire, M; Zanke, B W; Hudson, T J; Fertig, E J; Greenwood, C MT, and Hansen, K D. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biology*, 15(11):503, 2014. doi: 10.1186/s13059-014-0503-2.

Gebauer, K; Peters, I; Dubrowinskaja, N; Hennenlotter, J; Abbas, M; Scherer, R; Tezval, H; Merseburger, A S; Stenzl, A; Kuczyk, M A, and Serth, J. Hsa-mir-124-3 CpG island methylation is associated with advanced tumours and disease recurrence of patients with clear cell renal cell carcinoma. *British Journal of Cancer*, 108(1): 131–138, 2013. doi: 10.1038/bjc.2012.537.

Grandjean, V; Gounon, P; Wagner, N; Martin, L; Wagner, K D; Bernex, F; Cuzin, F, and Rassoulzadegan, M. The miR-124-Sox9 paramutation: RNA-mediated epigenetic control of embryonic and adult growth. *Development*, 136(21):3647–3655, 2009. doi: 10.1242/dev.041061.

Gribble, M O; Tang, W-Y; Shang, Y; Pollak, J; Umans, J G; Francesconi, K A; Goessler, W; Silbergeld, E K; Guallar, E; Cole, S A; Fallin, M D, and Navas-Acien, A. Differential methylation of the arsenic (III) methyltransferase promoter according to arsenic exposure. *Archives of Toxicology*, 88(2):275–282, 2014. doi: 10.1007/s00204-013-1146-x.

He, L; Vasiliou, K, and Nebert, D W. Analysis and update of the human solute carrier (SLC) gene superfamily. *Human Genomics*, 3(2):195–206, 2009.

IARC Working Group. Arsenic, metals, fibres, and dusts. *IARC Monographs on the Evaluation of Chronic Risks to Humans*, 100(Pt C):11–465, 2009.

Kaushal, A; Zhang, H; Karmaus, W J J; Everson, T M; Marsit, C J; Karagas, M R; Tsai, S-F; Wen, H-J, and Wang, S-L. Genome-wide DNA methylation at birth in relation to *in utero* arsenic exposure and the associated health in later life. *Environmental Health*, 16(1):50, 2017. doi: 10.1186/s12940-017-0262-0.

Kile, M L; Houseman, E A; Rodrigues, E; Smith, T J; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Su, L, and Christiani, D C. Toenail arsenic concentrations, GSTT1 gene polymorphisms, and arsenic exposure from drinking water. *Cancer Epidemiology Biomarkers and Prevention*, 14(10):2419–2426, 2005. doi: 10.1158/1055-9965.EPI-05-0306.

Kile, M L; Hoffman, E; Hsueh, Y-M; Afroz, S; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Ryan, L, and Christiani, D C. Variability in biomarkers of arsenic exposure and metabolism in adults over time. *Environmental Health Perspectives*, 117(3):455–460, 2009. doi: 10.1289/ehp.11251.

Kile, M L; Baccarelli, A; Hoffman, E; Tarantini, L; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Mostofa, G; Hsueh, Y-M; Wright, R O, and Christiani, D C. Prenatal arsenic exposure and DNA methylation in maternal and umbilical cord blood leukocytes. *Environmental Health Perspectives*, 120(7):1061–1066, 2012. doi: 10.1289/ehp.1104173.

Kile, M L; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Cardenas, A; Wright, R O, and Christiani, D C. Effect of prenatal

arsenic exposure on DNA methylation and leukocyte subpopulations in cord blood. *Epigenetics*, 9(5):774–82, 2014. doi: 10.4161/epi.28153.

Kile, M L; Cardenas, A; Rodrigues, E; Mazumdar, M; Dobson, C; Golam, M; Quamruzzaman, Q; Rahman, M, and Christiani, D C. Estimating effects of arsenic exposure during pregnancy on perinatal outcomes in a Bangladeshi cohort. *Epidemiology*, 27 (2):1, 2015. doi: 10.1097/EDE.0000000000000416.

Kitano, K; Watanabe, K; Emoto, N; Kage, H; Hamano, E; Nagase, T; Sano, A; Murakawa, T; Nakajima, J; Goto, A; Fukayama, M; Yatomi, Y; Ohishi, N, and Takai, D. CpG island methylation of microRNAs is associated with tumor size and recurrence of non-small-cell lung cancer. *Cancer Science*, 102(12):2126–2131, 2011. doi: 10.1111/j.1349-7006.2011.02101.x.

Küpers, L K; Xu, X; Jankipersadsing, S A; Vaez, A; la Bastide-van Gemert, S; Scholtens, S; Nolte, I M; Richmond, R C; Relton, C L; Felix, J F; Duijts, L; van Meurs, J B; Tiemeier, H; Jaddoe, V W; Wang, X; Corpeleijn, E, and Snieder, H. DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *International Journal of Epidemiology*, 44(4):1224–37, 2015. doi: 10.1093/ije/dyv048.

Liu, P; Piao, F-Y; Wang, Y-Y, and Hong, Y. Study on expression of neurotransmitter transmission - Related genes in cerebellum of mice exposed to arsenic by gene chip technology. *Journal of Dalian Medical University*, 30(5):401–04, 2008.

Liu, Y; Huang, X; Timani, K A; Broxmeyer, H E, and He, J J. MicroRNA-124 targets Tip110 expression and regulates hematopoiesis. *Stem Cells and Development*, 24(17): 2009–17, 2015. doi: 10.1089/scd.2015.0030.

Marsit, C J. Influence of environmental exposure on human epigenetic regulation. *Journal of Experimental Biology*, 218(1):71–79, 2015. doi: 10.1242/jeb.106971.

Naujokas, M F.; Anderson, B; Ahsan, H; Vasken Aposhian, H; Graziano, J H.; Thompson, C, and Suk, W A. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environmental Health Perspectives*, 121(3):295–302, 2013. doi: 10.1289/ehp.1205875.

Oberski, D. lavaan.survey: an R package for complex survey analysis of structural equation models. *Journal of Statistical Software*, 57(1):1–27, 2014. doi: 10.18637/jss.v057.i01. URL `http://www.jstatsoft.org/v57/i01/`.

Perera, F and Herbstman, J. Prenatal environmental exposures, epigenetics, and disease. *Reproductive Toxicology*, 31(3):363–373, 2011. doi: 10.1016/j.reprotox.2010.12.055.

Phookphan, P; Navasumrit, P; Waraprasit, S; Promvijit, J; Chaisatra, K; Ngaotepprutaram, T, and Ruchirawat, M. Hypomethylation of inflammatory genes (COX2,

EGR1, and SOCS3) and increased urinary 8-nitroguanine in arsenic-exposed newborns and children. *Toxicology and Applied Pharmacology*, 316:36–47, 2017. doi: 10.1016/j.taap.2016.12.015.

Platt, M J. Outcomes in preterm infants. *Public Health*, 128(5):399–403, 2014. doi: 10.1016/j.puhe.2014.03.010.

Quansah, R; Armah, F Ato; Essumang, D K; Luginaah, I; Clarke, E; Marfoh, K; Cobbina, S J; Nketiah-Amponsah, E; Namujju, P B; Obiri, S, and Dzodzomenyo, M. Association of arsenic with adverse pregnancy outcomes/infant mortality: a systematic review and meta-analysis. *Environmental Health Perspectives*, 123(5):412–421, 2015. doi: 10.1289/ehp.1307894.

Rahman, A; Vahter, M; Ekström, E-C, and Persson, L. Arsenic exposure in pregnancy increases the risk of lower respiratory tract infection and diarrhea during infancy in Bangladesh. *Environmental Health Perspectives*, 119(5):719–724, 2011. doi: 10.1289/ehp.1002265.

Reinius, L E; Acevedo, N; Joerink, M; Pershagen, G; Dahlén, S E; Greco, D; Söderhäll, C; Scheynius, A, and Kere, J. Differential DNA methylation in purified human blood cells: Implications for cell lineage and studies on disease susceptibility. *PLoS ONE*, 7(7), 2012. doi: 10.1371/journal.pone.0041361.

Rojas, D; Rager, J E; Smeester, L; Bailey, K A; Drobná, Z; Rubio-Andrade, M; Stýblo, M; García-Vargas, G, and Fry, R C. Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. *Toxicological Sciences*, 143(1):97–106, 2015. doi: 10.1093/toxsci/kfu210.

Saffery, R and Novakovic, B. Epigenetics as the mediator of fetal programming of adult onset disease: what is the evidence? *Acta Obstetricia et Gynecologica Scandinavica*, 93(11):1090–1098, 2014. doi: 10.1111/aogs.12431.

Schroeder, D I; Jayashankar, K; Douglas, K C; Thirkill, T L; York, D; Dickinson, P J; Williams, L E; Samollow, P B; Ross, P J; Bannasch, D L; Douglas, G C, and LaSalle, J M. Early developmental and evolutionary origins of gene body DNA methylation oatterns in mammalian placentas. *PLOS Genetics*, 11(8):e1005442, 2015. doi: 10.1371/journal.pgen.1005442.

Schroeder, J W; Conneely, K N; Cubells, J F; Kilaru, V; Newport, D J; Knight, B T; Stowe, Z N; Brennan, P A; Krushkal, J; Tylavsky, F A; Taylor, R N; Adkins, R M, and Smith, A K. Neonatal DNA methylation patterns associate with gestational age. *Epigenetics*, 6(12):1498–1504, 2011. doi: 10.4161/epi.6.12.18296.

Sitras, V; Paulssen, R; Leirvik, J; Vårtun, Å, and Acharya, G. Placental gene expression profile in intrauterine growth restriction due to placental insufficiency. *Reproductive Sciences*, 16(7):701–711, 2009. doi: 10.1177/1933719109334256.

Teschendorff, A E; Marabita, F; Lechner, M; Bartlett, T; Tegner, J; Gomez-Cabrero, D, and Beck, S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*, 29 (2):189–196, 2013. doi: 10.1093/bioinformatics/bts680.

Tsoporis, J N; Marks, A; Haddad, A; O'Hanlon, D; Jolly, S, and Parker, T G. S100A6 is a negative regulator of the induction of cardiac genes by trophic stimuli in cultured rat myocytes. *Experimental Cell Research*, 303(2):471–481, 2005. doi: 10.1016/j.yexcr.2004.10.018.

Tyler, C R and Allan, A M. The effects of arsenic exposure on neurological and cognitive dysfunction in human and rodent studies: a review. *Current Environmental Health Reports*, 1(2):132–147, 2014. doi: 10.1007/s40572-014-0012-1.

Xu, E; Gu, J; Hawk, E T; Wang, K K; Lai, M; Huang, M; Ajani, J, and Wu, X. Genome-wide methylation analysis shows similar patterns in Barrett's esophagus and esophageal adenocarcinoma. *Carcinogenesis*, 34(12):2750–2756, 2013. doi: 10.1093/carcin/bgt286.

Yu, J-Y; Chung, K-H; Deo, M; Thompson, R C, and Turner, D L. MicroRNA miR-124 regulates neurite outgrowth during neuronal differentiation. *Experimental Cell Research*, 314(14):2618–2633, 2008. doi: 10.1016/j.yexcr.2008.06.002.

6.8 Appendix

Table 6.7: Summary of associations between the CpGs associated with prenatal As exposure (EWAS) and birth gestational age. *P*-values and regression coefficients for unadjusted (reported) and cell-type adjusted models.

		Maternal water As, unadjusted for cell-type*			**Maternal water As, adjusted for estimated cell-types****			**Maternal postpartum toenail As, unadjusted for cell-type*****		
CpG ID	**Gene**	**β Coeff.**	***P***	***FDR***	**β Coeff.**	***P***	***FDR***	**β Coeff.**	***P***	***FDR***
cg01163597	SLC22A23	0.11	5.21×10^{-8}	5.44×10^{-4}	0.08	8.05×10^{-6}	0.03	0.08	1.75×10^{-2}	0.27
cg16081457		0.17	2.21×10^{-7}	5.81×10^{-4}	0.14	4.93×10^{-5}	0.04	0.12	4.57×10^{-2}	0.35
cg06522054	GNAL	0.11	1.64×10^{-7}	6.41×10^{-4}	0.1	8.50×10^{-6}	0.02	0.09	2.13×10^{-2}	0.28
cg20382695	ATRNL1	0.1	1.03×10^{-7}	5.76×10^{-4}	0.09	1.39×10^{-5}	0.03	0.07	8.24×10^{-2}	0.42
cg24937280	MCC	0.12	3.32×10^{-7}	7.07×10^{-4}	0.1	2.18×10^{-5}	0.04	0.09	4.52×10^{-2}	0.35
cg01910639	S100A6	-0.15	4.71×10^{-7}	7.70×10^{-4}	-0.09	6.21×10^{-4}	0.06	-0.09	1.09×10^{-1}	0.46
cg18115406	LMX1B	0.12	4.66×10^{-7}	7.68×10^{-4}	0.09	1.23×10^{-4}	0.04	0.06	1.55×10^{-1}	0.52
cg04874129	SLC6A2	0.14	9.00×10-7	8.67×10-4	0.1	1.22×10-4	0.04	0.1	6.45×10^{-2}	0.39
cg20277905	miR124-3	0.12	3.72×10^{-7}	7.27×10^{-4}	0.09	7.91×10^{-5}	0.04	0.12	7.32×10^{-3}	0.23
cg00398764		0.13	9.84×10-7	8.82×10-4	0.08	1.26×10-5	0.03	0.12	7.82×10^{-3}	0.23

*Model for $\log_2$(maternal drinking water As) adjusted for sex. **Model for $\log_2$(maternal drinking water As) adjusted for sex and estimated cell-type composition in cord blood (CD8T, CD4T, NK, B-cells, Monocytes, Granulocyte, and nucleated red blood cells). ***Model for $\log_2$(postpartum maternal toenail As) adjusted for sex.

Table 6.8: Relative contribution of each CpG to the main principal component (PC1) used in the discovery phase for mediation analyses.

CpG	**Gene**	**Relative loading on PC1**
cg01163597	SLC22A23	0.1
cg16081457		0.08
cg06522054	GNAL;GNAL;GNAL	0.1
cg20382695	ATRNL1	0.1
cg24937280	MCC	0.1
cg01910639	S100A6	0.1
cg18115406	LMX1B	0.11
cg04874129	SLC6A2	0.1
cg20277905	miR124-3	0.1
cg00398764		0.11

Table 6.9: CpG sites analyzed by bisulfite pyrosequencing.

Gene	**CpG**	**Genome coordinates**	**Mean methylation % (SD)**
miR124-3	CpG 22 (cg20277905)	Chr20:63178364	2.24 (2.68)
	CpG 21	Chr20:63178372	3.80 (3.95)
	CpG 20	Chr20:63178377	1.31 (2.06)
	CpG 19	Chr20:63178383	3.93 (4.07)
	CpG 18	Chr20:63178392	2.67 (2.96)
	CpG 17	Chr20:63178397	0.69 (1.92)
	CpG 16	Chr20:63178399	1.99 (2.14)
	CpG 15	Chr20:63178402	5.76 (2.68)
	CpG 14	Chr20:63178408	0.63 (1.56)
	CpG 13	Chr20:63178416	1.24 (3.23)
	CpG 12	Chr20:63178420	3.41 (2.97)
	CpG 11	Chr20:63178422	1.06 (1.93)
	CpG 10	Chr20:63178429	1.67 (4.11)
	CpG 9	Chr20:63178433	2.88 (3.78)
GNAL	CpG 1000	Chr18:11752462	1.05 (1.38)
	CpG 1001	Chr18:11752465	4.25 (2.24)
	CpG 1002	Chr18:11752478	2.14 (2.12)
	CpG 1003	Chr18:11752481	2.25 (2.13)
	CpG 1004	Chr18:11752494	2.37 (1.64)
	CpG 1005	Chr18:11752497	1.84 (1.41)
	CpG 1006	Chr18:11752499	1.27 (1.60)
	CpG 1007	Chr18:11752517	1.63 (1.80)
	CpG 1008 (cg06522054)	Chr18:11752539	3.19 (2.31)
MCC	CpG 48	Chr5:113488053	2.20 (2.67)
	CpG 47 (cg24937280)	Chr5:113488060	1.73 (2.26)
	CpG 46	Chr5:113488078	0.26 (1.06)
	CpG 45	Chr5:113488085	1.21 (1.76)
	CpG 44	Chr5:113488091	0.35 (1.08)
	CpG 43	Chr5:113488096	0.67 (1.28)
	CpG 42	Chr5:113488102	2.63 (2.16)
	CpG 41	Chr5:113488104	2.72 (2.08)
	CpG 40	Chr5:113488110	2.51 (2.14)
	CpG 39	Chr5:113488116	2.55 (2.17)
	CpG 38	Chr5:113488118	0.69 (1.87)
	CpG 37	Chr5:113488132	0.75 (1.13)
	CpG 36	Chr5:113488134	0.92 (1.30)

Table 6.10: Pearson correlations among CpG analyzed by bisulfite pyrosequencing located on *miR124-3*.

	CpG22 (cg20277905)	CpG21	CpG20	CpG19	CpG18	CpG17	CpG16	CpG15	CpG14	CpG13	CpG12	CpG11	CpG10
CpG21	0.20*†												
CpG20	0.34*†	0.17*†											
CpG19	0.10*†	0.09*†	0.14*†										
CpG18	0.24*†	0.17*†	0.27*†	0.20*†									
CpG17	0.16*†	0.07	0.26*†	0.13*†	0.12*†								
CpG16	0.19*†	0.18*†	0.25*†	0.19*†	0.28*†	0.19*†							
CpG15	0.33*†	0.3	0.28*†	0.09*†	0.19*†	0.08	0.22*†						
CpG14	0.22*†	0.17*†	0.14*†	0.09*†	0.12*†	0.16*†	0.23*†	0.30*†					
CpG13	0.08	0.05	0.10*†	0.06	0.09*†	0.16*†	0.13*†	0.01	0.14*†				
CpG12	0.36*†	0.25*†	0.29*†	0.24*†	0.27*†	0.22*†	0.40*†	0.42*†	0.29*†	0.13*†			
CpG11	0.27*†	0.13*†	0.35*†	0.17*†	0.22*†	0.27*†	0.28*†	0.26*†	0.35*†	0.20*†	0.33*†		
CpG10	0.12*†	0.16*†	0.21*†	0.08	0.16*†	0.16*†	0.24*†	0.21*†	0.11*†	0.14*†	0.23*†	0.29*†	
CpG9	0.10*†	0.14*†	0.15*†	0.16*†	0.17*†	0.16*†	0.20*†	0.11*†	0.08	0.46*†	0.26*†	0.16*†	0.12*†

*P <0.05, unadjusted; †P<0.05, FDR adjusted.

Table 6.11: Pearson correlations among CpG analyzed by bisulfite pyrosequencing located on *GNAL*.

	CpG1000	CpG1001	CpG1002	CpG1003	CpG1004	CpG1005	CpG1006	CpG1007
CpG1001	0.34*†							
CpG1002	0.25*†	0.41*†						
CpG1003	0.33*†	0.37*	0.49*					
CpG1004	0.34*†	0.57*†	0.67*†	0.48*†				
CpG1005	0.28*†	0.34*†	0.30*†	0.41*†	0.59*†			
CpG1006	0.44*†	0.44*†	0.31*†	0.50*†	0.49*†	0.54*†		
CpG1007	0.21*†	0.25*†	0.26*†	0.33*†	0.52*†	0.70*†	0.45*†	
CpG1008 (cg06522054)	0.22*†	0.34*†	0.25*†	0.36*†	0.52*†	0.67*†	0.42*†	0.68*†

*P <0.05, unadjusted; †P<0.05, FDR adjusted.

Table 6.12: Pearson correlations among CpG analyzed by bisulfite pyrosequencing located on *MCC*.

	CpG36	CpG37	CpG38	CpG39	CpG40	CpG41	CpG42	CpG43	CpG44	CpG45	CpG46	CpG47 (cg24937280)
CpG37	0.42*†											
CpG38	0.31*†	0.25*†										
CpG39	0.53*†	0.41*†	0.27*†									
CpG40	0.45*†	0.39*†	0.26*†	0.71*†								
CpG41	0.41*†	0.36*†	0.23*†	0.49*†	0.41*†							
CpG42	0.44*†	0.38*†	0.26*†	0.55*†	0.46*†	0.61*†						
CpG43	0.49*†	0.39*†	0.30*†	0.45*†	0.46*†	0.38*†	0.46*†					
CpG44	0.40*†	0.33*†	0.20*†	0.29*†	0.22*†	0.27*†	0.34*†	0.31*†				
CpG45	0.43*†	0.52*†	0.29*†	0.37*†	0.36*†	0.44*†	0.49*†	0.47*†	0.33*†			
CpG46	0.27*†	0.27*†	0.18*†	0.19*†	0.28*†	0.33*†	0.37*†	0.33*†	0.32*†	0.31*†		
CpG47 (cg24937280)	0.46*†	0.28*†	0.29*†	0.43*†	0.49*†	0.47*†	0.50*†	0.42*†	0.21*†	0.38*†	0.41*†	
CpG48	0.40*†	0.33*†	0.22*†	0.33*†	0.35*†	0.34*†	0.42*†	0.34*†	0.25*†	0.38*†	0.36*†	0.56*†

*P <0.05, unadjusted; †P<0.05, FDR adjusted.

Table 6.13: Parameter estimates from the structural equation model for the mediated effect of maternal drinking water As concentration on birth gestational age through variation of DNA methylation of *miR124-3* in the validation phase.

Pathway	**Effect**	**β Coeff. (95% CIs)**	***P***
$\log_2$(Water As) → *miR124-3*	Direct	0.05 (0.01, 0.10)	0.03
$\log_2$(Water As) → Gestational age	Direct	-0.14 (-0.20, -0.09)	<0.001
miR124-3 → Gestational age	Direct	-0.41 (-0.60, -0.22)	<0.001
$\log_2$(Water As) → *miR124-3* → Gestational age	Indirect	-0.02 (-0.04, 0.00)	0.03
SEM fit measures (robust): $\chi^2 = 140.908$, $P = 0.006$; CFI = 0.943; RMSEA = 0.032.			

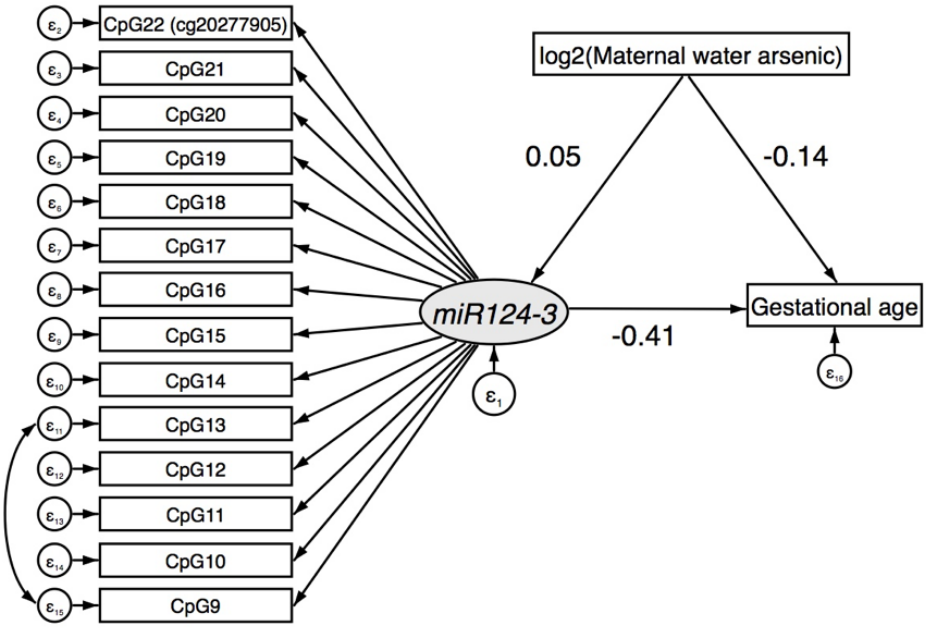


Figure 6.11: Structural equation model for the mediated association of maternal drinking water As concentration and birth gestational age by DNA methylation of *miR124-3* in the validation phase.

Table 6.14: Parameter estimates from the structural equation model for the mediated effect of maternal drinking water As concentration on birth gestational age through variation of DNA methylation of *GNAL* in the validation phase.

Pathway	**Effect**	**β Coeff. (95% CIs)**	***P***
$\log_2$(Water As) → *GNAL*	Direct	0.03 (0.01, 0.05)	0.02
$\log_2$(Water As) → Gestational age	Direct	-0.16 (-0.21, -0.10)	<0.001
GNAL → Gestational age	Direct	-0.28 (-0.60, 0.02)	0.07
$\log_2$(Water As) → *GNAL* → Gestational age	Indirect	-0.01 (-0.02, 0.00)	0.17
SEM fit measures (robust): $\chi^2 = 80.483$, P <0.001; CFI = 0.953; RMSEA = 0.069.			

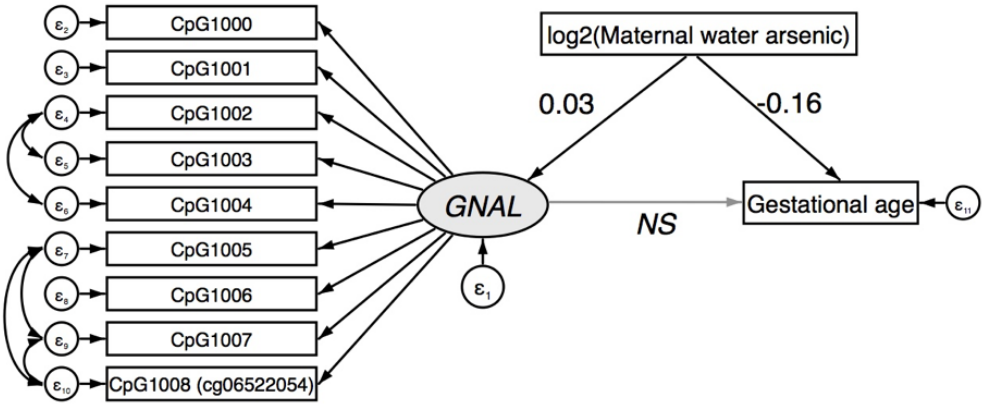


Figure 6.12: Structural equation model for the mediated association of maternal drinking water As concentration and birth gestational age by DNA methylation of *GNAL* in the validation phase. Abbreviation: NS, non-significant.

Table 6.15: Parameter estimates from the structural equation model for the mediated effect of maternal drinking water As concentration on birth gestational age through variation of DNA methylation of *MCC* in the validation phase.

Pathway	**Effect**	**β Coeff. (95% CIs)**	***P***
$\log_2$(Water As) → *MCC*	Direct	0.06 (0.03, 0.09)	0.001
SNP rs1057827 → *MCC* (reference: A/A)			
G/G	Direct	0.52 (0.07, 0.97)	0.03
A/G	Direct	0.13 (-0.05, 0.31)	0.17
$\log_2$(Water As) → Gestational age	Direct	-0.14 (-0.20, -0.08)	<0.001
MCC → Gestational age	Direct	-0.47 (-0.70, -0.24)	<0.001
$\log_2$(Water As) → *MCC* → Gestational age	Indirect	-0.03 (-0.05, -0.01)	0.1004
SEM fit measures (robust): $\chi^2 = 207.179$, P <0.001; CFI = 0.943; RMSEA = 0.048.			

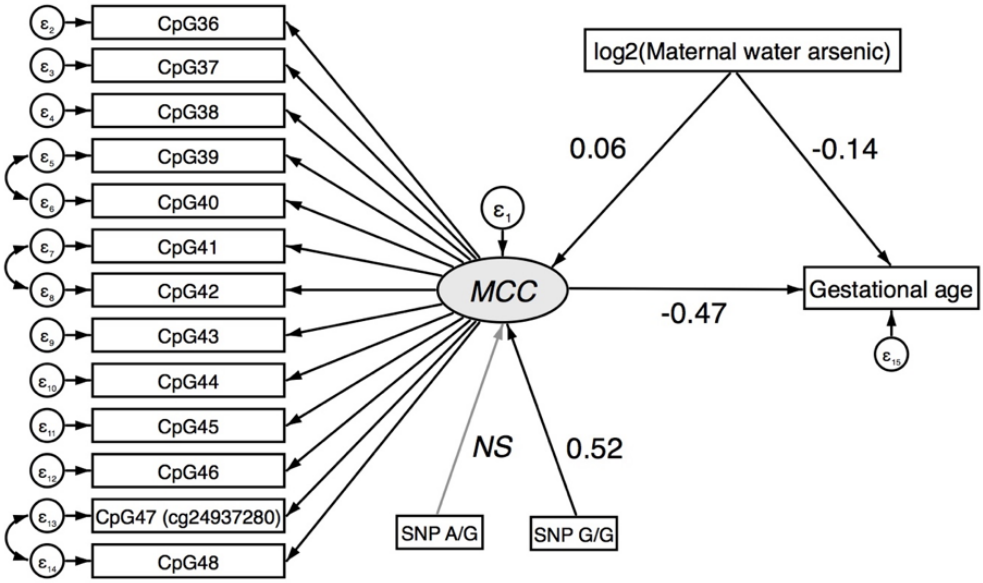


Figure 6.13: Structural equation model for the mediated association of maternal drinking water As concentration and birth gestational age by DNA methylation of *MCC* in the validation phase. Abbreviation: NS, non-significant.

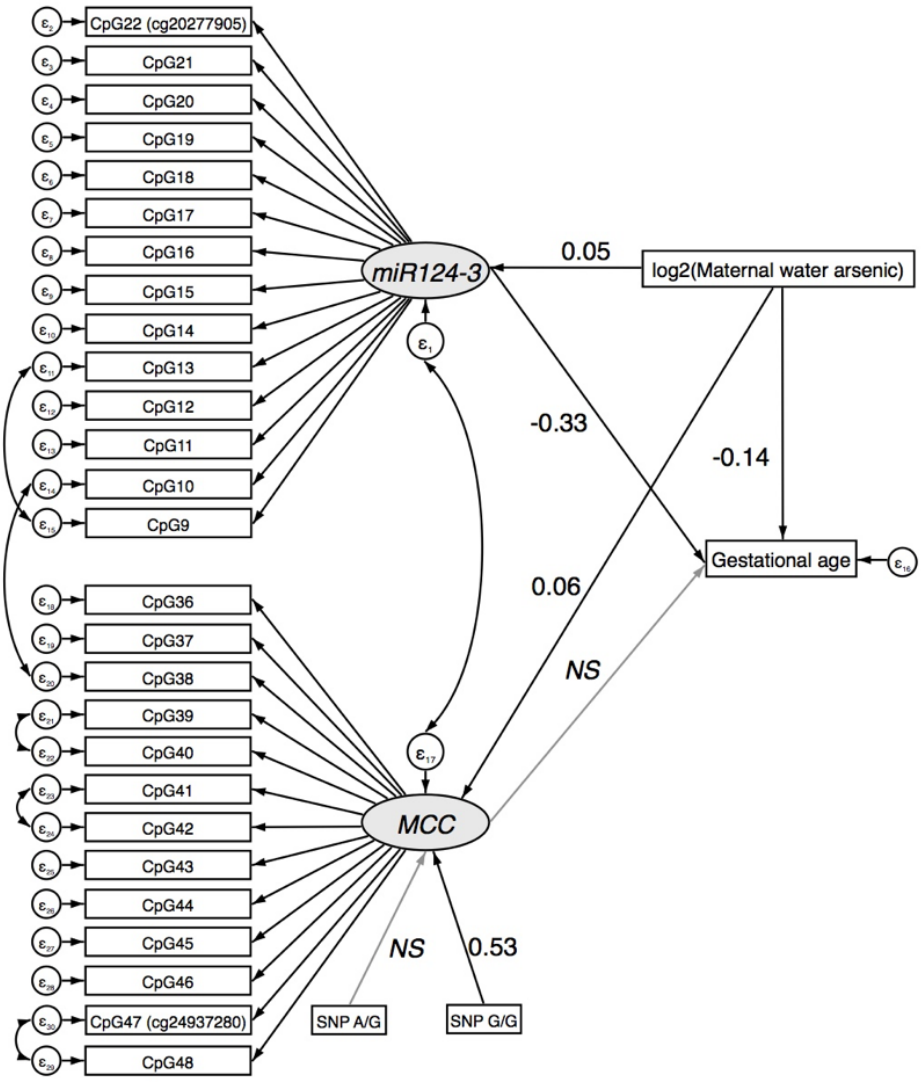


Figure 6.14: Unadjusted structural equation model for the mediated effect of maternal drinking water As concentration on birth gestational age through variation of DNA methylation of CpGs in *miR124-3* and *MCC* in the validation phase. Abbreviation: NS, non-significant.

Table 6.16: Unadjusted structural equation model for the mediated effect of maternal drinking water As concentration on birth gestational age through variation of DNA methylation of CpGs in *miR124-3* and *MCC* in the validation phase.

Pathway	**Effect**	**β Coeff. (95% CIs)**	***P***
miR124-3			
$\log_2$(Water As) → *miR124-3*	Direct	0.05 (0.01, 0.10)	0.03
MCC			
$\log_2$(Water As) → *MCC*	Direct	0.06 (0.03, 0.09)	<0.001
SNP rs1057827 → *MCC* (reference: AA)			
G/G	Direct	0.53 (0.12, 0.94)	0.01
A/G	Direct	0.11 $(-0.05, 0.27)$	0.19
Gestational age			
$\log_2$(Water As) → Gestational age	Direct	$-0.14(-0.20, -0.08)$	<0.001
miR124-3 → Gestational age	Direct	$-0.33(-0.53, -0.12)$	0.002
MCC → Gestational age	Direct	$-0.18(-0.47, 0.10)$	0.2
$\log_2$(Water As) → *miR124-3* → Gestational age	Indirect	$-0.02(-0.03, 0.00)$	0.05
$\log_2$(Water As) → *MCC* → Gestational age	Indirect	$-0.01(-0.03, 0.01)$	0.22
SEM fit measures (robust): $\chi^2 = 718.542$, P <0.001; CFI = 0.899; RMSEA = 0.041.			

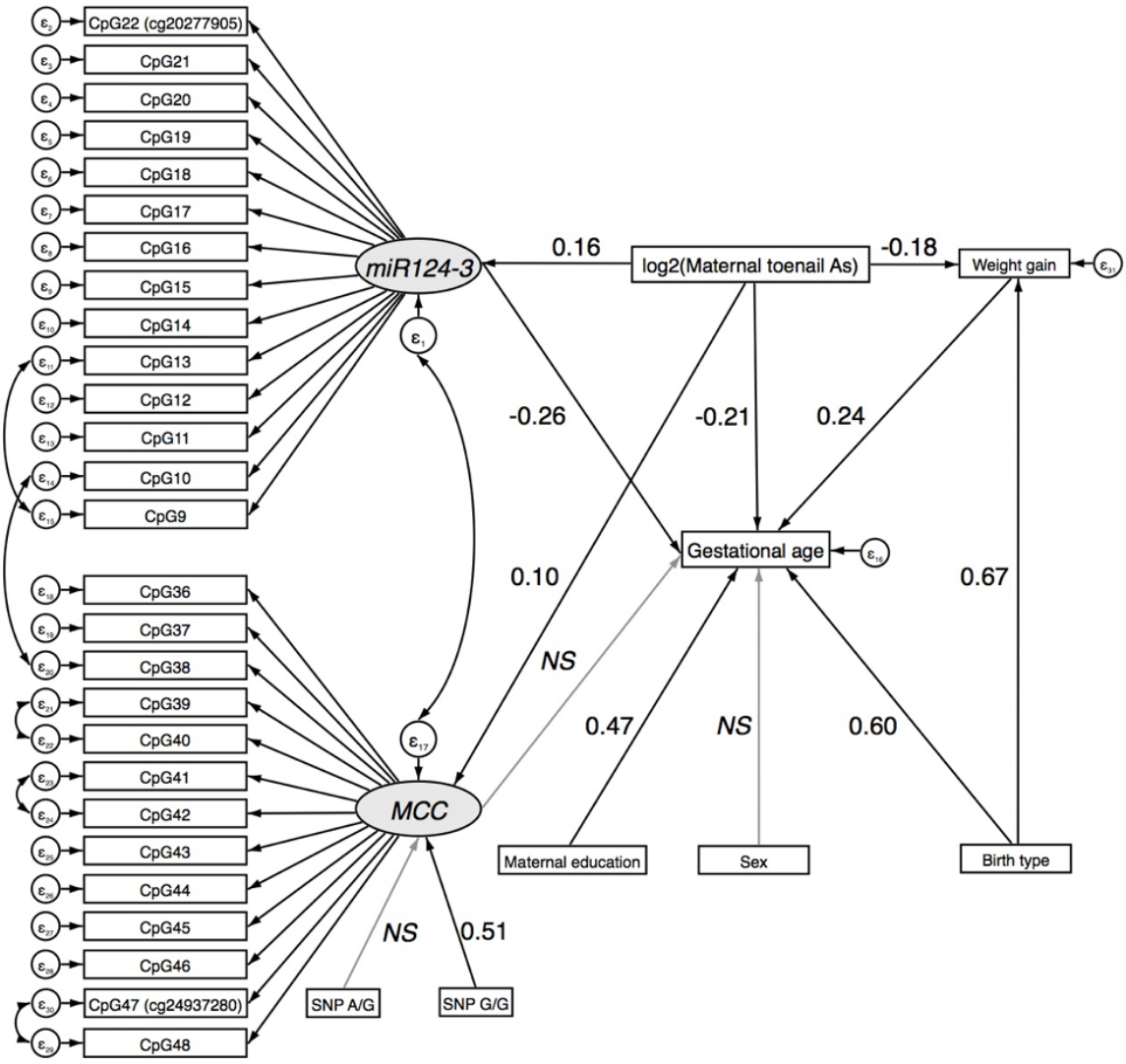


Figure 6.15: Adjusted structural equation model for the mediated effect of postpartum maternal toenail As concentration on birth gestational age through variation of DNA methylation of CpGs in *miR124-3* and *MCC* in the validation phase. Abbreviation: NS, non-significant.

Table 6.17: Structural equation model for the mediated effect of postpartum maternal toenail As concentration on birth gestational age through variation of DNA methylation of CpGs in *miR124-3* and *MCC* in the validation phase.

Pathway	**Effect**	**β Coeff. (95% CIs)**	***P***
miR124-3			
$\log_2$(Maternal toenail As) → *miR124-3*	Direct	0.16 (0.07, 0.25)	0.001
MCC			
$\log_2$(Maternal toenail As) → *MCC*	Direct	0.10 (0.04, 0.15)	0.001
SNP rs1057827 (reference: AA)			
G/G	Direct	0.51 (0.10, 0.92)	0.01
A/G	Direct	0.10 (−0.07, 0.25)	0.26
Maternal weight gain			
$\log_2$(Maternal toenail As) → Weight gain	Direct	-0.18 (−0.33, −0.03)	0.02
Birth type (Cesarean vs. vaginal) → Weight gain	Direct	0.67 (0.15, 1.19)	0.01
Gestational age			
$\log_2$(Maternal toenail As) → Gestational age	Direct	−0.21(−0.34, −0.08)	0.002
miR124-3 → Gestational age	Direct	−0.26(−0.46, −0.06)	0.01
MCC → Gestational age	Direct	-0.18 (-0.45, 0.09)	0.2
Infant ex (Male vs. female) → Gestational age	Direct	0.11 (-0.20, 0.42)	0.48
Weight gain → Gestational age	Direct	0.24 (0.18, 0.30)	<0.001
Birth type (Cesarean section vs. vaginal) → Gestational age	Direct	0.60 (0.30, 0.90)	<0.001
Maternal education (> primary vs. ≤ primary) → Gestational age	Direct	0.47 (0.15, 0.79)	0.004
$\log_2$(Maternal toenail As) → Weight gain → Gestational age	Indirect	−0.04(−0.08, −0.01)	0.03
$\log_2$(Maternal toenail As) → *miR124-3* → Gestational age	Indirect	−0.04(−0.08, −0.01)	0.02
$\log_2$(Maternal toenail As) → *MCC* → Gestational age	Indirect	-0.02 (-0.04, 0.01)	0.22
SEM fit measures (robust): $\chi^2 = 886.434$, P <0.001; CFI = 0.893; RMSEA = 0.038.			

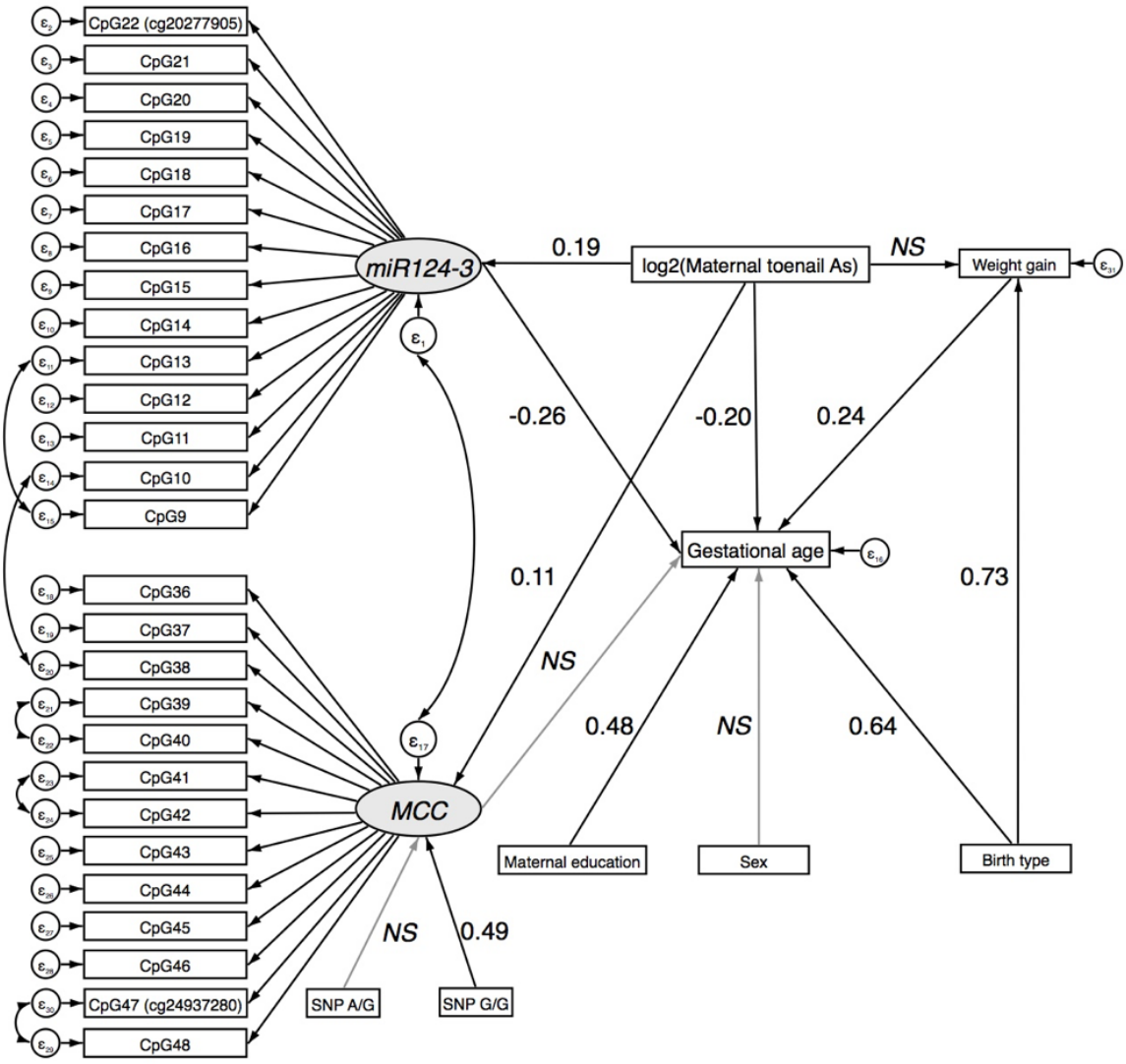


Figure 6.16: Adjusted structural equation model for the mediated effect of maternal toenail As concentration at enrollment on birth gestational age through variation of DNA methylation of CpGs in *miR124-3* and *MCC* in the validation phase. Abbreviation: NS, non-significant.

Table 6.18: Structural equation model for the mediated effect of maternal toenail As concentration at enrollment on birth gestational age through variation of DNA methylation of CpGs in *miR124-3* and *MCC* in the validation phase.

Pathway	**Effect**	**β Coeff. (95% CIs)**	***P***
miR124-3			
$\log_2$(Maternal toenail As) → *miR124-3*	Direct	0.19 (0.09, 0.29)	<0.001
MCC			
$\log_2$(Maternal toenail As) → *MCC*	Direct	0.11 (0.05, 0.17)	<0.001
SNP rs1057827 → *MCC* (reference: AA)			
G/G	Direct	0.49 (0.08, 0.90)	0.02
A/G	Direct	0.10 (-0.07, 0.26)	0.25
Maternal weight gain			
$\log_2$(Maternal toenail As) → Weight gain	Direct	-0.11 $(-0.28, 0.05)$	0.18
Birth type (Cesarean vs. vaginal) → Weight gain	Direct	0.73 (0.22, 1.25)	0.01
Gestational age			
$\log_($Maternal toenail As) → Gestational age	Direct	-0.20 $(-0.34, -0.07)$	0.003
miR124-3 → Gestational age	Direct	$-0.26(-0.46, -0.06)$	0.01
MCC → Gestational age	Direct	$-0.18(-0.45, 0.10)$	0.2
Infant sex (Male vs. female) → Gestational age	Direct	0.09 $(-0.22, 0.40)$	0.56
Weight gain → Gestational age	Direct	0.24 (0.18, 0.30)	<0.001
Birth type (Cesarean section vs. vaginal) → Gestational age	Direct	0.64 (0.34, 0.94)	<0.001
Maternal education (> primary vs. ≤ primary) → Gestational age	Direct	0.48 (0.16, 0.80)	0.003
$\log_($Maternal toenail As) → Weight gain → Gestational age	Indirect	$-0.03(-0.07, 0.01)$	0.18
$\log_2$(Maternal toenail As) → $miR124-3$ → Gestational age	Indirect	$-0.05(-0.09, 0.01)$	0.02
$\log_2$(Maternal toenail As) → *MCC* → Gestational age	Indirect	$-0.02(-0.05, 0.01)$	0.21
SEM fit measures (robust): χ^2= 884.11 4, P <0.001; CFI = 0.894; RMSEA = 0.038.			

Chapter 7

DNA methylation of *DNMT3A* mediates the association between *in utero* arsenic exposure and birth outcomes: results from a prospective birth cohort in Bangladesh

Anne K. Bozack,[1] Andres Cardenas,[2] Quazi Quamruzzaman,[3] Mahmuder Rahman,[3], Golam Mostofa,[3] David C. Christiani,[4], Molly L. Kile[5]

Affiliations: 1. Dept. of Environmental Health Sciences, Columbia University; 2. Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley; 3. Dhaka Community Hospital, Dhaka, Bangladesh; 4. Dept. of Environmental Health, Harvard T.H. Chan School of Public Health; 5. School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University

7.1 Abstract

Background: Epigenetic programming plays a critical role in fetal development. DNA methyltransferase 3 alpha (DNMT3A), an enzyme responsible for *de novo* DNA methylation, is a prime candidate gene as a mediator between prenatal exposures and birth outcomes. *In utero* arsenic (As) exposure has been associated with reduced birth weight and gestational age. We evaluated the relationships between *in utero* As exposure, birth outcomes, and *DNMT3A* DNA methylation.

Methods: In a prospective Bangladeshi birth cohort, cord blood DNA methylation of three *DNMT3A* CpGs was measured using bisulfite pyrosequencing. Maternal toenail As concentrations at birth were measured to estimate *in utero* exposure. Among vaginal births (N = 413), structural equation models (SEMs) were used to evaluate relationships between *DNMT3A* methylation, $\log_2$(maternal toenail As concentration), birth weight, and gestational age.

Results: In an SEM including birth weight and gestational age, and adjusted for infant sex, maternal weight gain, and maternal education, maternal toenail As levels were associated with *DNMT3A* DNA methylation and gestational age ($P < 0.05$). *DNMT3A* methylation was associated with gestational age and birth weight ($P < 0.05$). There was a significant indirect effect of As on gestational age through *DNMT3A* ($P = 0.024$) and a borderline significant indirect effect on birth weight ($P = 0.082$). However, there were significant indirect effects of maternal toenail As levels on birth weight through pathways including gestational age ($P = 0.038$), *DNMT3A* methylation and gestational age ($P = 0.037$), and maternal weight gain and gestational age ($P = 0.012$). The total effect of a doubling in maternal toenail As concentration is a decrease in gestational age of 2.1 days and a decrease in birth weight of 28.9 grams ($P = 0.001$).

Conclusions: DNMT3A plays a critical role in fetal epigenetic programming.

In utero As exposure was associated with more methylation of CpG sites in *DNMT3A* which appeared to mediate the association between prenatal As exposure and birth outcomes. Additional studies are needed to verify this finding.

7.2 Introduction

Chronic exposure to arsenic (As) persists in many regions of the world. In Bangladesh, approximately 40 million individuals rely on household drinking water with As concentrations exceeding the World Health Organization (WHO) guideline of 10 μg/L, half of which are also above the Bangladesh standard 50 μg/L (Bangladesh Bureau of Statistics and United Nation Children's Fund, 2015). Inorganic As and As metabolites readily pass the placenta, resulting in a high correlation between As concentrations measured in maternal and cord blood (Concha et al., 1998; Hall et al., 2007). Arsenic is an established human toxicant and group 1 carcinogen (World Health Organization, 2011) and maternal As exposure during fetal development has been linked to increased risk of adverse health outcomes later in life including cancers of the lung, bladder, liver, and larynx, cardiovascular disease, and reduced lung function (Farzan et al., 2013; Vahter, 2009). *In utero* and early life exposure has been associated with increased childhood morbidity (Farzan et al., 2013; Rahman et al., 2017).

The teratogenic effects of As have been well established in rodent models (Nagymajtényi et al., 1985; Ferm and Hanlon, 1985; Hill et al., 2008; Kozul-Horvath et al., 2012; Hood, 1972; Hood and Bishop, 1972; Morrissey and Mottet, 1983; Moore et al., 2019). However, findings from epidemiological studies of the association between *in utero* As exposure and birth outcomes in humans have been inconsistent, possibly due to differences in study design, level of exposure, exposure assessment, and sample size. Considering the cumulative evidence, a recent review by Milton et al. found an insufficient number of studies addressing neonatal death and preterm birth but consistent and convincing evidence of the positive association As exposure and risk of spontaneous abortion, stillbirth, and low birth weight (Milton et al., 2017). Additionally, a meta-analysis conducted by Zhong et al. found a negative association between *in utero* As exposure and birth weight (summary regression coefficient from 12 studies

= -25.0 g; 95% CI: -41.0, -9.0) (Zhong et al., 2019). Although numerous epidemiological studies have addressed the association between *in utero* As exposure and birth weight, decreased birth weight is the result of shortened gestation and/or intrauterine grown restriction, factors with varying effects on health outcomes (Wardlaw et al., 2004). Shortened gestation is associated with increased risk of infant mortality, morbidity, and disability, whereas restricted intrauterine growth is associated with decreased growth in childhood and increased morbidity in later in life. Therefore, mediation analyses including birth weight and gestational age are important to understanding the health effects of *in utero* As exposure. Using structural equation models (SEMs), in a birth cohort in Bangladesh, it was found that the negative association between *in utero* As exposure and birth weight is fully mediated by decreased gestational age (β = -17.37, 95% CI: -22.77, -11.98) and maternal weight gain (β = -1.80, 95% CI: -3.72, 0.13) (Kile et al., 2015).

Changes in the epigenome, including DNA methylation, may be one mechanism underlying the associations between As exposure and health outcomes (Bailey et al., 2016; Bjørklund et al., 2018). As explained by the developmental origins of health and disease hypothesis, environmental exposures during embryogenesis, a time of cellular differentiation and epigenetic reprogramming, may result in epigenetic dysregulation and increased disease risk (Heindel and Vandenberg, 2015). The DNA methyltransferases 3 alpha (DNMT3A) and 3 beta (DNMT3B) are responsible for *de novo* methylation in embryonic cells (Okano et al., 1999). DNMT3A and DNMT3B have differential spatial and temporal expression during early embryogenesis (Uysal et al., 2017). DNMT3A may have a more significant role in maintaining global DNA methylation, particularly at distal promotors and non-CpG sites (Gu et al., 2018), and in imprinting (Kaneda et al., 2004). Due to the critical role of DNMT3A in establishing *de novo* DNA methylation, we hypothesized that DNA methylation of this gene could

mediate the association between As exposure occurring in the critical prenatal development phase and adverse birth outcomes, specifically birth weight and gestational age. We also hypothesized the gestational age would act as a mediator between *in utero* As exposure, epigenetic changes, and birth weight. We explored this hypothesis using SEMs with our *a priori* hypothesis depicted in (Figure 7.1).

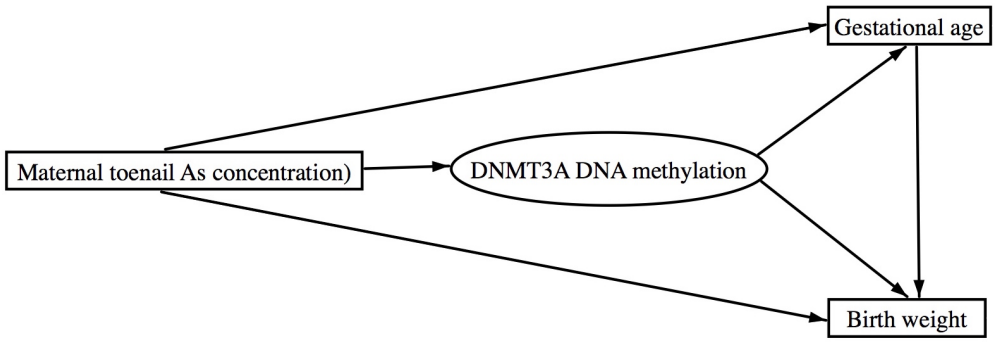


Figure 7.1: Conceptual structural equation model for the mediated association of *in utero* arsenic exposure and birth outcomes by DNA methylation of *DNMT3A*

7.3 Methods

Study Population

Participants were recruited as part or prospective birth cohort of women exposed to As through drinking water in Bangladesh in the Sirajdikhan and Pabna Sadar upazilas. This cohort has been described in detail previously (Kile et al., 2014). Women were recruited at $\leq$ 16 weeks gestational age by Dhaka Community Hospital (DCH) Trust (N = 1,458). Eligibility criteria were having a single pregnancy, using a tube well as the main source of drinking water, and planning to remain in the current residence for the duration of the pregnancy. Women received free prenatal care provided by DCH, as well as prenatal vitamins, which were given to participants during monthly

home visits. Women gave birth with DCH-trained medical personnel at a local clinic or at their home; cord blood was collected at birth. A sub-sample of 569 newborns from the cohort were selected for measurement of DNA methylation in whole cord blood using pyrosequencing. This sub-sample was randomly selected across a range of arsenic exposure that was measured in the households drinking water. We also restricted analyses to only vaginal births to minimize potential confounding which limited the final sample size to 413 newborns.

Ethics

The study protocol was approved by the Human Research Committees at the Harvard School of Public Health, Oregon State University, and DCH Trust. Prior to enrollment in the study, informed consent was obtained from all participants.

Laboratory measures

Maternal toenail As and quality control: Arsenic concentrations in maternal toenail samples were measured as a biomarker of exposure during pregnancy. Arsenic has an affinity for sulfhydryl groups and accumulates in scleroproteins, and therefore toenail As concentration represents levels of ingested As during nail growth (Karagas et al., 1996). Toenail clippings were collected from mothers at the time of enrollment and within one month of delivery. The methods used to measure arsenic exposure have been previously published (Rodrigues et al., 2015). Briefly, samples were sonicated in 1% Triton X-100 solution (Sigma-Aldrich, Inc., St. Louis, MO) and rinsed in Milli-Q water (Millipore Corporation, Billerica, MA) to remove external contamination. Then nails were digested in Trace Select Ultra Pure nitric acid (HNO_3; Sigma-Aldrich, Inc.), and diluted with Milli-Q water. Inductively coupled plasma mass spectrometry was used to measure total arsenic concentration (Perkin-Elmer Model DRC-II 6100, Nor-

walk, CT). Arsenic concentrations of human hair references were used to correct for method error using black correction and normalization (CRM Hair; Shanghai Institute of Nuclear Research, Academia Sinica, China); human hair was selected as the reference due to lack of available toenail references. Eighteen postpartum toenail samples were missing. Samples were dropped from analyses if they had a mass < 5 mg ($N = 1$) or relative standard deviation $> 25\%$ ($N = 3$). One sample below the batch limit of detection (range: 0.004 - 0.85 μg As/g) was also excluded.

Maternal drinking water As and quality control: At the time of enrollment, participants were asked to identify their main source of drinking water (Kile et al., 2014). Water samples were collected, preserved with nitric acid to a pH < 2, and stored at room temperature. Arsenic concentrations in water samples were measured using inductively coupled mass spectrometry (ICP-MS) with the US EPA method 200.8 (Environmental Laboratory Services, North Syracuse, NY) (Creed et al., 1994). Samples had an average percent recovery from plasmaCal multi-element QC standard #1 solution (SCP Science) of $102 \pm 7\%$. Thirty samples below the LOD of 1 μg/L were replaced with LOD/2.

Bisulfite pyrosequencing: DNA methylation was quantified for cg26544247 (GRCh37/hg19, chr2:25,473,782) located on the north shore of a CpG island within the gene body of *DNMT3A*. Coverage also included two downstream CpG sites (chr2: 25,473,813 and chr2:25,473,843). Bisulfite pyrosequencing was performed at EpigenDx (Hopkinton, MA) using 20 ng/ul whole cord blood DNA. Bisulfite conversion was used to convert unmethylated cytosines to uracil. Following PCR amplification and direct pyrosequencing, the average percent DNA methylation was calculated. Control samples with low methylation, medium methylation, and high methylation were included on each plate.

Other variables: Sociodemographic data were collected by trained interviewers during clinical visits. Infant birth weight was measured using a pediatric scale

calibrated before each use. Gestational age in weeks was determined by ultrasound measurements taken at enrollment.

Statistical analysis

Descriptive statistics were calculated for all variables of interest and covariates (mean and SD for continuous variables, frequency for categorical variables). Maternal toenail As concentrations were right skewed and therefore $\log_2$ transformed. We assessed relationships between loci with quantified DNA methylation (cg26544247, chr2:25473813, and chr2:25473843) using Pearson correlations. Mediation was evaluated using structural equation models (SEMs) that included DNA methylation of all measured loci. SEMs are a multivariate statistical technique that allow for confirmatory analysis of a given hypothesis (Gunzler et al., 2013). By simultaneously estimating multiple and related regression-like models, SEMs allow for a given variable to act as both an independent and dependent variable (the term exogenous is used to refer to variables that act only as independent variables, and the term endogenous is used to refer to variables that as depended variables in at least one modeled relationship). SEMs use underlying latent (or unmeasured) variables defined by observed variables representing same construct, therefore addressing multicollinearity. DNA methylation of cg26544247, chr2:25473813, and chr2:25473843 was used as indicator variables to construct a latent variable representing DNA methylation of *DNMT3A*.

Mediation of the association between $\log_2$(postpartum maternal toenail As concentrations) and the outcomes of interest (birthweight and gestational age) by the latent variable was assessed in a single SEM model that was adjusted for the potential confounders of infant sex, maternal weight gain between enrollment and delivery, and maternal education ($>$ primary vs. $\leq$ primary education). Birth weight was modeled in kg to ensure variances were similarly scaled. The full information maximum likeli-

hood approach (FIML) was used to estimate model parameters, and robust estimates of model fit were used due to skewness in DNAm variables (Brosseau-Liard et al., 2012; Brosseau-Liard and Savalei, 2014). For models including categorical covariates (i.e., infant sex and maternal education), standard errors were calculated from 10,000 bootstrap samples. In each SEM, model indices were examined to determine if including residual correlations would improve model fit.

Our primary measure of exposure was postpartum maternal toenail As concentrations. We conducted sensitivity analyses using $\log_2$(maternal toenail As concentrations at enrollment) and $\log_2$(maternal water As concentrations) as the measures of exposure. In addition, we used linear models to test the associations between DNA methylation of the target CpG site (cg26544247), $\log_2$(postpartum maternal toenail As concentrations), birth weight, and gestational age. Specifically, we evaluated associations between the exposure (maternal toenail As concentration) and outcome (gestational age and birthweight), the mediator (DNA methylation) and the exposure (maternal toenail As concentration), and the mediator (DNA methylation) and the outcome (gestational age and birthweight).

All statistical analyses were performed using the R statistical package, version 3.5.0 (R Core Team, 2015). SEMs were conducted using the R *lavaan* package (Oberski, 2014).

7.4 Results

Participant Characteristics

Measures of DNA methylation was available for 413 infants with vaginal births. Maternal and infant characteristics are summarized in Table 1. Approximately half of the infants were male (52.7%). The median ± IQR maternal drinking water As

Table 7.1: Participant characteristics (N = 413)

	Median ± IQR	**Range**
Drinking water As at recruitment (μg/L)	17 ± 73.5	0.5 - 545.0
Maternal toenail As at enrollment	2.1 ± 3.7	0.1 - 51.3
Postpartum maternal toenail As (ng/μg)[a]	1.7 ± 3.2	0.04 - 34.8
Gestational age at delivery (weeks)[b]	38 ± 3	22 - 42
Preterm birth (< 37 weeks gestation), N (%)	120 (29.0%)	
Birth weight (g)[c]	2,800 ± 600	1,400 - 4,600
Low birth weight (< 2,500 g), N (%)	107 (25.8%)	
Infant sex		
Male, N (%)	218 (52.7%)	
Female, N (%)	196 (47.3%)	
Percent methylation		
cg26544247 (chr2: 25,473,782)[c]	2.5 ± 6.1	0.0 - 25.0
chr2:25,473,813	2.1 ± 5.3	0.0 - 19.7
chr2:25,473,843	0 ± 3.6	0.0 - 25.4

a. N = 390; b. N = 411; c. GRCh37/hg19. Abbreviation: IQR,interquartile range.

concentration at recruitment was 17.0 ± 73.5 μg/L (range: 0.5 - 545.0 μg/L), and the mean maternal toenail As concentration at delivery was 1.7 ± 3.2 ng/μg (range: 0.04 - 34.8 ng/μg). The median gestational age was 38 ± 3 weeks (range: 22 - 42 weeks), and the median birth weight was 2,800 ± 600 g (range: 1,400 - 4,600 g). Consistent with other birth cohort studies in Bangladesh, there was a high rate of preterm birth (< 37 weeks gestation, 29.0%) (Shah et al., 2014) and low birth weight (< 2,500 g, 25.9%) (Monawar Hosain et al., 2005).

There were low levels of DNA methylation at each of the CpG sites. The median ± IQR percent methylation for the CpG sites were 2.5 ± 6.1, 2.1 ± 5.3, 0.0 ± 3.6 for cg26544247 (chr2:25,473,782), chr2:25,473,813, and chr2:25,473,843, respectively. DNA methylation of was significantly and positively correlated between the three CpG sites measured (cg26544247 and chr2:25,473,813: $r_{Pearson} = 0.93$, $P < 0.001$, cg26544247 and chr2:25,473,843: $r_{Pearson} = 0.92$, $P < 0.001$, chr2:25,473,813 and chr2:25,473,843: $r_{Pearson} = 0.86$, $P < 0.001$) (Figure 7.2). There was a trend toward increasing DNA methylation with As exposure at each CpG site. Figure 7.3 shows the median and IQR

of DNA methylation by quartiles of postpartum maternal toenail As concentration.

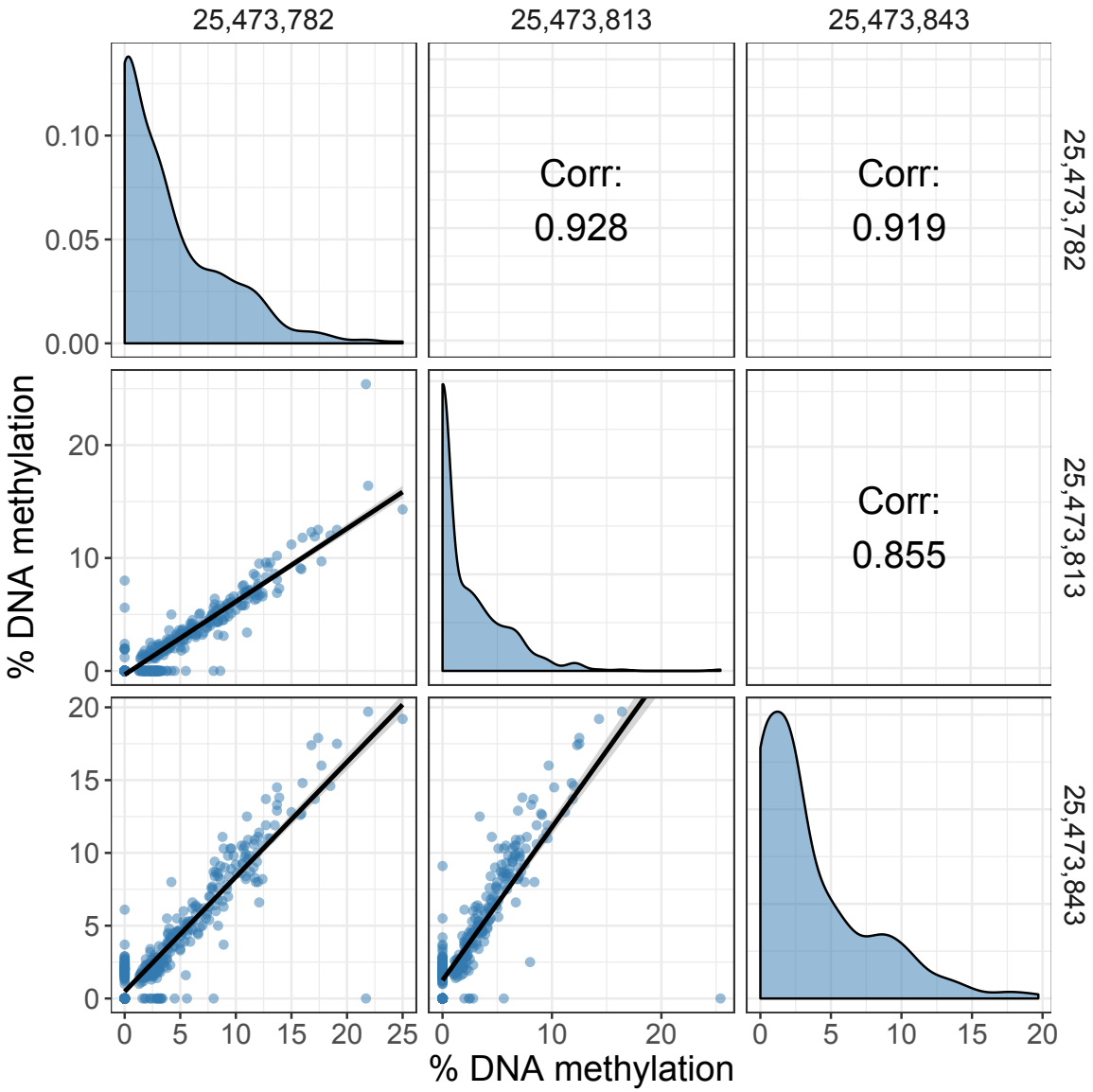


Figure 7.2: Correlation matrix of DNA methylation of CpGs measured in *DNMT3A*. Correlations are presented in the top right cells, scatter plots of % DNA methylation are presented in the bottom left cells, and density plots of % DNA methylation are presented in the diagonal cells.

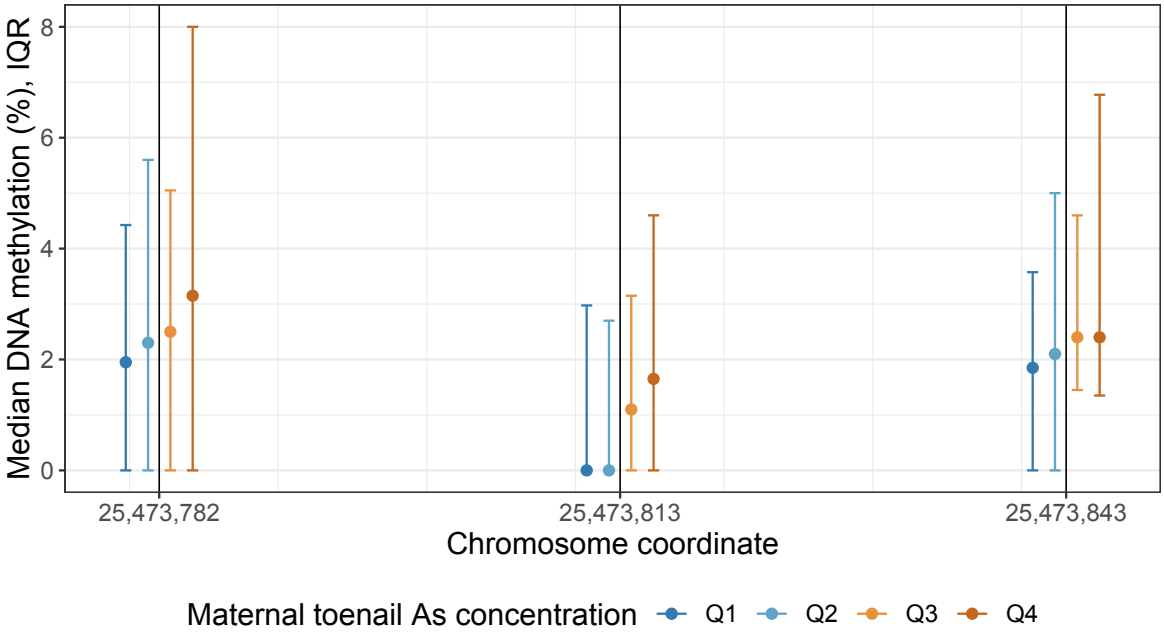


Figure 7.3: Median and IQR of *DNMT3A* DNA methylation by quartiles of postpartum maternal toenail As concentration. Q1 $\leq$ 0.84 ng/μg, Q2 $>$ 0.84 ng/μg and $\leq$ 1.66 ng/μg, Q3 $>$ 1.66 ng/μg and $\leq$ 4.00 ng/μg, Q4 $>$ 4.00 ng/μg.

Mediation analysis

We first assessed mediation of the associations between $\log_2$(postpartum maternal toenail As concentrations) and birth outcomes by DNA methylation of *DNMT3A* in separate SEMs for birth weight and gestational age. A latent variable representing *DNMT3A* methylation was constructed using percent methylation of cg26544247, chr2:25,473,813, and chr2:25,473,843. We observed a significant positive association between maternal As levels and DNMT3A DNA methylation (B = 0.37, P = 0.004) in an SEM for the mediated effect of $\log_2$(postpartum maternal toenail As concentrations) on birth weight through *DNMT3A* DNA methylation. We also observed a significant positive association between maternal As levels and *DNMT3A* DNA methylation (B = 0.37, P = 0.004), and a significant negative association between *DNMT3A* DNA methylation and birth weight (B = -20.67 g, P $<$ 0.001) (Appendix Table 7.6 and

Appendix Figure 7.6). The direct effect of maternal As levels on birth weight was not significant (B = 1.62 g, P = 0.91); however, the indirect effect of maternal As levels on birth weight through *DNMT3A* DNA methylation was statistically significant (B = -7.62 g, P = 0.015). Similarly, in an SEM for the mediated effect of $\log_2$(postpartum maternal toenail As concentrations) on gestational age through *DNMT3A* methylation, there was a significant positive association between maternal As levels and *DNMT3A* DNA methylation (B = 0.39, P = 0.002), and a significant negative association between *DNMT3A* DNA methylation and gestational age (B = -0.12 weeks, $P < 0.001$) (Appendix Table 7.7 and Appendix Figure 7.7). Both the direct and indirect effects of maternal As levels on gestational age were statistically significant (direct effect: B = -0.25 weeks, P = 0.004; indirect effect: B = -0.05 weeks, P = 0.018).

Subsequently, we constructed a single SEM to test for the indirect effects of $\log_2$(postpartum maternal toenail As concentrations) on birth weight and gestational age through *DNMT3A* DNA methylation. This model accounted for an indirect effect of maternal As exposure on birth weight though gestational age, but due to the results of individual models, we did not include the direct effect of As exposure on birth weight (Tables 7.2 and 7.3, and Figure 7.4). As observed in the individual models, there was a significant positive association between $\log_2$(postpartum maternal toenail As concentrations) and *DNMT3A* methylation (B = 0.39, P = 0.002), and negative associations between maternal toenail As levels and gestational age (B = -0.25 weeks, P = 0.004), and between *DNMT3A* DNA methylation and gestational age (B = -0.12 weeks, $P < 0.001$). There were also significant indirect effects of maternal toenail As levels on birth weight through *DNMT3A* methylation (B = -0.05 weeks, P = 0.017). The direct effect of *DNMT3A* DNA methylation on birth weight and the indirect effect of maternal toenail As levels on birth weight through *DNMT3A* DNA methylation were borderline significant (B = -10.12 g, P = 0.074; B = -3.98 g, P = 0.088, respectively).

Table 7.2: Unadjusted SEM for the mediated effect of postpartum maternal toenail As levels on gestational through *DNMT3A* DNA methylation, and the mediated effect of postpartum maternal toenail As levels on birth weight through gestational age and *DNMT3A* DNA methylation.

Pathway	**Effect**	***B***	**(95% CI)**	***P***
DNMT3A				
$\log_2$(Maternal toenail As) → *DNMT3A* DNAm	Direct	0.39	(0.14, 0.65)	0.002
Gestational age				
$\log_2$(Maternal toenail As) → Gestational age	Direct	-0.25	(-0.42, -0.08)	0.004
DNMT3A DNAm → Gestational age	Direct	-0.12	(-0.18, -0.06)	< 0.001
log(Maternal toenail As) → *DNMT3A* DNAm → Gestational age	Indirect	-0.05	(-0.09, -0.01)	0.017
Birth weight				
DNMT3A DNAm → Birth weight	Direct	-10.12	(-21.23, 0.99)	0.074
Gestational age → Birth weight	Direct	77.37	(62.01, 92.73)	< 0.001
log(Maternal toenail As) → *DNMT3A* DNAm → Birth weight	Indirect	-3.98	(-8.55, 0.59)	0.088
$\log_2$(Maternal toenail As) → gestational age → Birth weight	Indirect	-19.52	(-33.50, -5.52)	0.006
$\log_2$(Maternal toenail As) → *DNMT3A* DNAm → Gestational age → Birth weight	Indirect	-3.6	(-6.77, -0.44)	0.026

Abbreviatio: DNAm, DNA methylation

Table 7.3: Fit indices for the unadjusted SEM for the mediated effect of postpartum maternal toenail As levels on gestational through *DNMT3A* DNA methylation, and the mediated effect of postpartum maternal toenail As levels on birth weight through gestational age and *DNMT3A* DNA methylation.

Index	**Criterion for good fit**	**Model fit[1]**
χ^2 *P*-value	>0.05	0.245
Root Mean Square Error of Approximation (RMSEA)	< 0.05	0.025
Comparative Fit Index (CFI)	>0.95	0.999
Tucker-Lewisnon-normed Fit Index	>0.90	0.998
Standardized Root Mean Squared Residual	>0.05	0.012

1. Fit indices calculated from robust estimators.

However, we observed a significant association between gestational age and birth weight ($B = 77.37$ g, $P < 0.001$) and significant indirect effects of maternal toenail As levels on birth weight through gestational age ($B =$ -19.52 g, $P = 0.006$) and through *DNMT3A* DNA methylation and gestational age ($B =$ -3.60 g, $P = 0.026$).

Similar results were observed in an SEM adjusted for infant sex, maternal weight gain, and maternal education (Table 7.4, and Figure 7.5). Fit indices for this final model indicated good model fit (Table 7.5). Maternal toenail As levels were

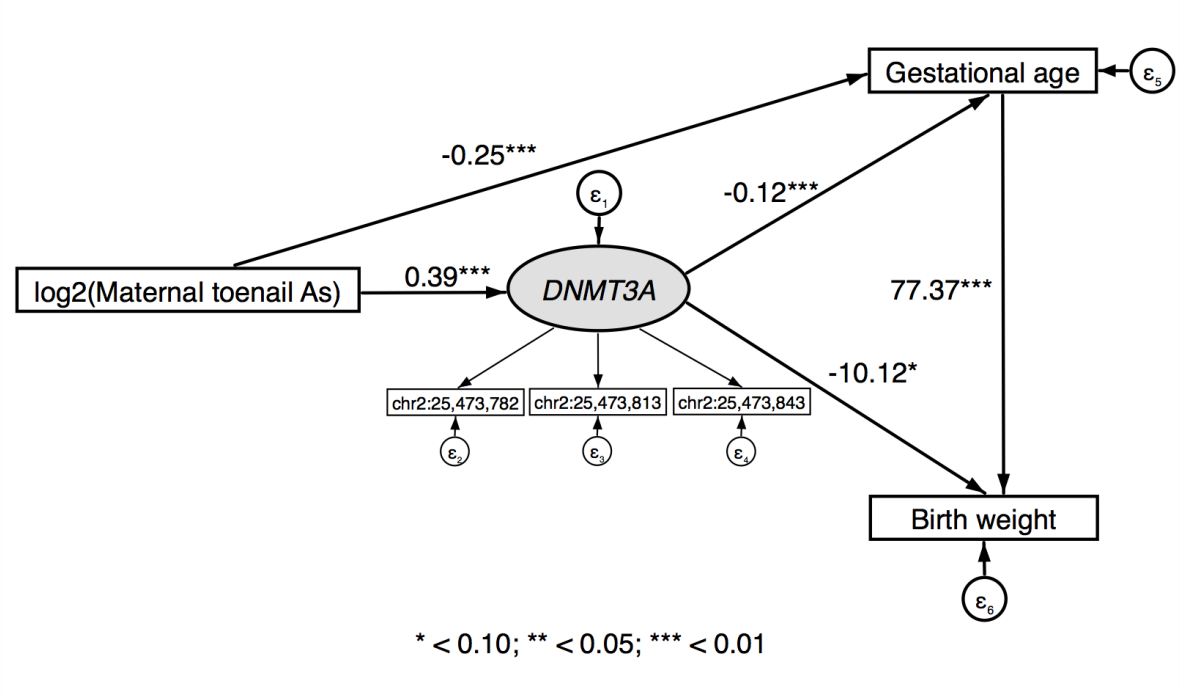


Figure 7.4: Unadjusted SEM for the mediated effect of postpartum maternal toenail As levels on gestational through *DNMT3A* DNA methylation, and the mediated effect of postpartum maternal toenail As levels on birth weight through gestational age and *DNMT3A* DNA methylation.

significantly associated with *DNMT3A* DNA methylation (B = 0.40, P = 0.002) and gestational age (B = -0.19 weeks, P = 0.004), and *DNMT3A* DNA methylation was significantly associated with gestational age (B = -0.19 weeks, P = 0.027) and birth weight (B = -11.03 g, P = 0.048). There was a significant indirect effect of maternal toenail As levels on gestational age through *DNMT3A* methylation (B = -0.04 weeks, P = 0.024), however the indirect effect on birth weight remained borderline significant (B = -4.37 g, P = 0.082). There were significant indirect effects of maternal toenail As levels on birth weight through gestational age (B = -14.44 g, P = 0.038), through *DNMT3A* DNA methylation and gestational age (B = -3.14 g, P = 0.037), and through maternal weight gain and gestational age (B = -5.14 g, P = 0.012). Calculated from the adjusted SEM, the total effect of a doubling in maternal toenail As concentration is a decrease in gestational age of 2.07 days (P = 0.001) and a decrease in birth weight

of 28.89 g ($P = 0.001$).

7.5 Sensitivity analyses

Our primary measure of exposure was postpartum maternal toenail As concentrations. Sensitivity analyses were performed using $\log_2$(maternal toenail As concentrations at enrollment) and $\log_2$(maternal drinking water arsenic concentrations) as the measures of exposure. Postpartum maternal toenail As concentrations were significantly correlated with maternal toenail As concentrations at enrollment ($r_{Spearman} = 0.80$, $P < 0.001$) and maternal drinking water As concentrations ($r_{Spearman} = 0.61$, $P < 0.001$). Results from adjusted SEMs using maternal toenail As levels collected postpartum and at enrollment were consistent: maternal toenail As levels at enrollment were associated with *DNMT3A* DNA methylation ($B = 0.40$, $P = 0.003$) and gestational age (B = -0.24 weeks, $P = 0.007$), *DNMT3A* DNA methylation was associated with gestational age (B = -0.10 weeks, $P = 0.001$) and birth weight (B = -11.01 g, $P = 0.053$), and there were significant indirect effects of maternal toenail As levels on birth weight through gestational age (B = -18.25 g, $P = 0.01$) and through *DNMT3A* DNA methylation and gestational age (B = -3.11 g, $P = 0.045$) (Appendix Table 7.8 and Figure 7.8). However, in an adjusted SEM using maternal drinking water As levels, we observed a smaller effect size of the association between As exposure and *DNMT3* DNA methylation ($B = 0.18$, $P = 0.004$). The association between maternal drinking water As levels and gestational age, and the indirect effect of maternal drinking water As levels on birth weight through gestational age were not significant, although the effect estimates were negative (B = -0.06 weeks, $P = 0.12$; B = -4.52 g, $P = 0.13$, respectively) (Appendix Table 7.9 and Figure 7.9).

We also assessed the associations between *in utero* As exposure, birth outcomes, and *DNMT3* DNA methylation using linear models to confirm results obtained

Table 7.4: Adjusted SEM for the mediated effect of postpartum maternal toenail As levels on gestational through *DNMT3A* DNA methylation, and the mediated effect of postpartum maternal toenail As levels on birth weight through gestational age and *DNMT3A* DNA methylation, controlling for infant sex, maternal weight gain, and maternal education.

Pathway	**Effect**	***B***	**(95% CI)**	***P***
DNMT3A				
$\log_2$(Maternal toenail As) → *DNMT3A* DNAm	Direct	0.4	(0.15, 0.66)	0.002
Infant sex → *DNMT3A* DNAm	Direct	-0.75	(-1.47, -0.04)	0.041
Maternal weight gain				
$\log_2$(Maternal toenail As) → Maternal weight gain	Direct	-0.24	(-0.41, -0.07)	0.004
Gestational age				
$\log_2$(Maternal toenail As) → Gestational age	Direct	-0.19	(-0.36, -0.03)	0.027
DNMT3A DNAm → Gestational age	Direct	-0.1	(-0.16, -0.04)	0.001
Maternal weight gain → Gestational age	Direct	0.28	(0.20, 0.36)	<0.001
Maternal education (> primary vs. ≤ primary) → Gestational age	Direct	0.56	(0.17, 0.97)	0.006
$\log_2$(Maternal toenail As) → *DNMT3A* DNAm → Gestational age	Indirect	-0.04	(-0.08, -0.01)	0.024
$\log_2$(Maternal toenail As) → Maternal weight gain → Gestational age	Indirect	-0.07	(-0.12 -0.02)	0.006
Birth weight				
DNMT3A DNAm → Birth weight	Direct	-11.03	(-21.54, 0.41)	0.048
Gestational age → Birth weight	Direct	77.04	(58.08, 94.39)	<0.001
Infant sex → Birth weight	Direct	-118.96	(-196.45, -41.25)	0.003
Maternal weight gain → Birth weight	Direct	7.46	(-10.60, 26.00)	0.425
Maternal education → Birth weight	Direct	-60.82	(-140.15, 20.83)	0.139
log(Maternal toenail As) → *DNMT3A* DNAm → Birth weight	Indirect	-4.37	(-9.69, 0.17)	0.081
log(Maternal toenail As) → Gestational age → Birth weight	Indirect	-14.44	(-29.24, -1.92)	0.038
log(Maternal toenail As) → *DNMT3A* DNAm → Gestational age → Birth weight	Indirect	-3.14	(-6.59, -0.77)	0.037
log(Maternal toenail As) → Maternal weight gain → Birth weight	Indirect	-1.8	(-6.73, 2.82)	0.44
$\log_2$(Maternal toenail As) → Maternal weight gain → Gestational age → Birth weight	Indirect	-5.14	(-9.60, -1.48)	0.012

Abbreviation: DNAm, DNA methylation.

Table 7.5: Fit indices for the adjusted SEM for the mediated effect of postpartum maternal toenail As levels on gestational through *DNMT3A* DNA methylation, and the mediated effect of postpartum maternal toenail As levels on birth weight through gestational age and *DNMT3A* DNA methylation, controlling for infant sex, maternal weight gain, and maternal education.

Index	**Criterion for good fit**	**Model fit$_1$**
χ^2 *P*-value	>0.05	0.166
Root Mean Square Error of Approximation (RMSEA)	< 0.05	0.028
Comparative Fit Index (CFI)	>0.95	0.997
Tucker-Lewisnon-normed Fit Index	>0.90	0.994
Standardized Root Mean Squared Residual	>0.05	0.028

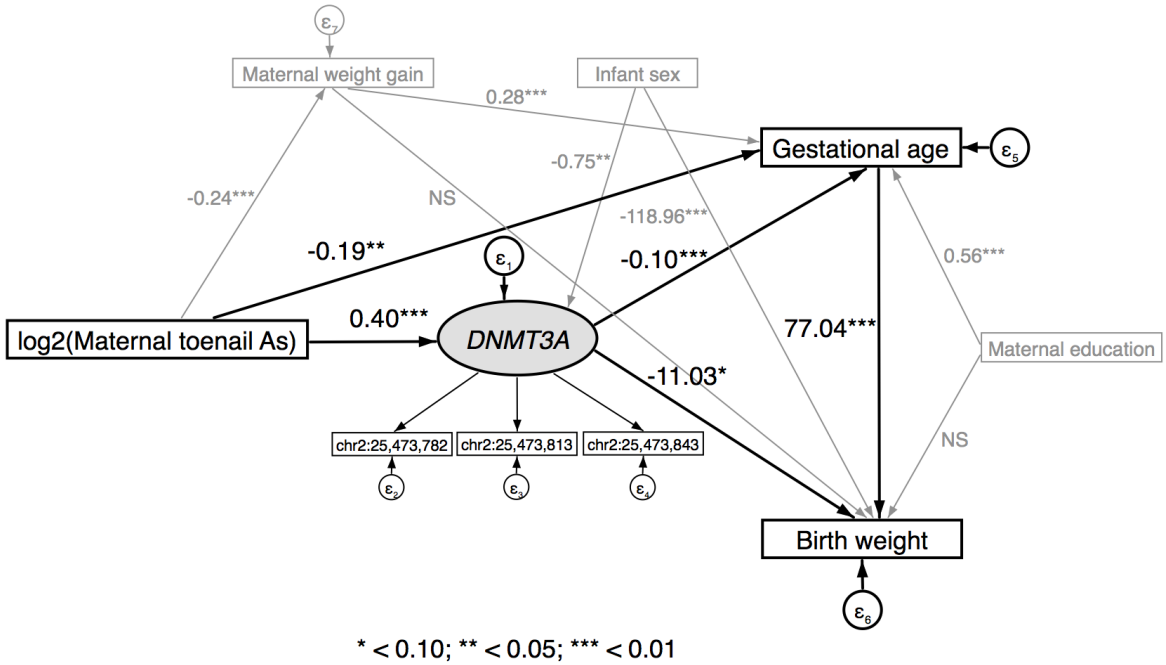


Figure 7.5: Adjusted SEM for the mediated effect of postpartum maternal toenail As levels on gestational through *DNMT3A* DNA methylation, and the mediated effect of postpartum maternal toenail As levels on birth weight through gestational age and *DNMT3A* DNA methylation, controlling for infant sex, maternal weight gain, and maternal education.

through SEMs. Specifically, we evaluated associations between the exposure (postpartum maternal toenail As levels) and outcomes (birth weight and gestational age), the mediator (represented by the target CpG site *DNMT3*, cg26544247) and the exposure (postpartum maternal toenail As levels), and the mediator (*DNMT3* DNA methyla-

tion) and the outcomes (birthweight and gestational age). Results from unadjusted models and models adjusted for infant sex, maternal weight gain, and maternal education are shown in Appendix Table 7.10. Results from crude and adjusted linear models were consistent with SEMs. In adjusted models, $\log_2$(postpartum maternal toenail As concentrations) were associated with decreased gestational age (B = -0.16 weeks, P = 0.010) and increased methylation of cg26544247 (B = 0.44, P = 0.005), DNA methylation of cg26544247 was associated with decreased gestational age (B = -0.09 weeks, P < 0.001) and birth weight (B = -12.6 g, P < 0.001), and gestational age was associated with birth weight (B = 79.1 g, P < 0.001). Maternal As levels were not significantly associated with birth weight (B = 1.9 g, P = 0.89).

7.6 Discussion

Using an SEM approach, we assessed the association between *in utero* As exposure and birth outcomes, and mediation by DNA methylation of *DNMT3A*. SEMs are advantageous for mediation analysis involving variables located on a causal pathway (i.e., variables that act as a dependent variable in one model and as an independent variable in another model) (Gunzler et al., 2013). Furthermore, SEMs incorporate latent variables representing multiple, related indicator variables (i.e., measured variables) addressing limitations posed by multicollinearity in traditional biostatistical approaches (Maruyama, 1998). Although commonly used in other fields of research, SEMs can be a valuable tool in environmental epidemiology for understanding complex relationships (Buncher et al., 1991). In this study, SEMs allowed us to create a latent variable representing DNA methylation of three correlated CpG sites located in the gene body of *DNMT3A* and simultaneously evaluate the direct and indirect effects of *in utero* As exposure on birth outcomes.

We observed that DNA methylation of *DNMT3A* mediated the associa-

tion between *in utero* As exposure and birth outcomes. We found that increased $\log_2$(postpartum maternal toenail As concentrations) were associated with increased *DNMT3A* methylation and decreased gestational age, and there was a significant negative indirect effect of As on gestational age through *DNMT3A* methylation. In addition, we observed significant negative indirect effects of maternal toenail As levels on birth weight through gestational age and though *DNMT3A* methylation and gestational age. The indirect effect of maternal As levels on birth weight through *DNMT3A* methylation alone was borderline significant. Overall, a doubling in maternal toenail As concentration was associated with a decrease in gestational age of 2.1 days and a decrease in birth weight of 28.9 g.

Our results are consistent with other studies finding a negative association between *in utero* As exposure and gestational age (Xu et al., 2011; Röllin et al., 2017), although some studies have reported a borderline significant negative association (Laine et al., 2015) or a null association (Sun et al., 2019; Freire et al., 2019; Bloom et al., 2016). It should be noted that epidemiological studies have generally analyzed associations with preterm birth ($<$ 37 weeks gestation), rather than gestational age (Bloom et al., 2014). However, health effects later in life may be associated with early term, in addition to preterm, birth (Boyle et al., 2012). Furthermore, this clinical categorization may not be sufficient to detect variation in gestational age, and may be not be appropriate for populations with high rates of preterm birth such as Bangladesh (Shah et al., 2014). Although we did not observe a direct effect of *in utero* As exposure on birth weight, we did find a significant total effect of As exposure on birth weight fully mediated through pathways including gestational age; mediation through DNA methylation of *DNMT3A* alone was borderline significant. Multiple epidemiological studies have investigated the effects of *in utero* As exposure on birth weight (as reviewed by Bloom et al. (2014); Milton et al. (2017); Zhong et al. (2019)). However,

our results indicate the importance of including gestational age as a birth outcome of interest, and addressing relationships between birth outcomes. Furthermore, decreased birth weight may be due to shortened gestational period and/or intrauterine grown restriction (Wardlaw et al., 2004). Investigating the etiology of observed reductions in birth weight is important to understanding the health effects of *in utero* As exposure.

Consistent with our findings, DNA methylation of cg26544247, the target site, has previously been negatively associated with gestational age in a Norwegian birth cohort ($FDR < 0.05$) (Bohlin et al., 2016). This study found differential methylation throughout the epigenome (44,359 probes at $FDR < 0.05$ for ultrasound-estimated gestational age and 44,544 probes for last menstrual period-estimate gestational age at $FDR < 0.05$). However, 18 additional CpG sites annotated to *DNMT3A* were identified as differentially methylated and a gene ontology analysis identified pathways including *DNMT3A*. In addition, in birth cohort in Tennessee, one CpG annotated to *DNMT3A* was identified as associated with gestational age ($FDR < 0.05$) (Parets et al., 2013). Methylation at additional CpG sites located in *DNM3A* have been associated with the *in utero* environment. Specifically, increased methylation in cord blood of cg13344237, located within 200 base pairs of the transcription start site, has been associated with lower maternal pre-pregnancy BMI (underweight < 18.5 kg/m^2 vs. normal 18.5 - 24.9 kg/m^2) ($FDR < 0.05$) (Sharp et al., 2015). Increased methylation in cord blood of cg15843262, located in the gene body, was associated with antidepressant use during pregnancy ($P = 8\text{x}10^{-4}$) (Non et al., 2014).

The role of DNMT3A in *de novo* DNA methylation is well-established (Okano et al., 1999), and DNMT3A is expressed during embryogenesis (Watanabe et al., 2002). However, how DNMT3A may affect birth outcomes is not well understood. DNMT3A is involved in multiple, diverse biological pathways; Gene Ontology (GO) terms annotated to *DNMT3A* include roles in DNA methylation, RNA polymerase binding and

transcription, genetic imprinting, mitotic cell cycle, protein binding, aging, and regulation of cell death. In addition, it is not known how DNA methylation at our target CpG site, cg26544247, is associated with gene expression. Further research is needed to understand the biological pathways involved in the observed association between *DNMT3A* methylation and birth outcomes.

Our study is strengthened by the prospective birth cohort design with multiple measures of As exposure during pregnancy. Toenail As concentrations reflect exposure during the prior several months to a year (Kile et al., 2005). Maternal toenail samples were collected within one moth of delivery to provide an estimate of internal dose during pregnancy and establish temporality. Drinking water samples and maternal toenail samples were also collected at the time of enrollment early in pregnancy. Overall, results from sensitivity analyses using maternal drinking water As levels and maternal toenail As levels at enrollment as the measures of exposure were consistent. In addition, gestational age was determined by ultrasound at the time of enrollment, which provides a more accurate estimation than reported last menstrual period.

Several limitations of our study should be noted. We did have missing data on maternal toenail As concentrations (N = 21 for maternal toenail concentrations at enrollment, N = 23 for postpartum maternal toenail concentrations). Gestational age and birth weight were lower among mother-infant pairs with missing maternal toenail As concentrations at enrollment (mean gestational age: with missing data = 34.3 weeks, without missing data = 37.6 weeks; Kruskal-Wallis $P = 0.003$; mean birth weight: with missing data = 2,481 g, without missing data = 2,764 g; Kruskal-Wallis $P = 0.012$) and postpartum (mean gestational age: with missing data = 34.5 weeks, without missing data = 37.6 weeks; Kruskal-Wallis $P = 0.005$; mean birth weight: with missing data = 2,510 g, without missing data = 2,763 g; Kruskal-Wallis $P = 0.018$). However, we had full data on maternal water As concentrations, and maternal drinking water

concentrations were significantly correlated and toenail As concentrations at enrollment ($r_{Spearman} = 0.60$, $P < 0.001$) and postpartum ($r_{Spearman} = 0.61$, $P < 0.001$). Although in sensitivity analyses using maternal water As levels we observed reduced significance of the indirect effect of *in utero* As on birth weight through *DNMT3* methylation and gestational age, the confidence intervals were wide and tended toward a negative association ($\log_2$(maternal water As) $\rightarrow$ *DNMT3A* DNAm $\rightarrow$ birth weight: B (95% CI) = -2.01 (-4.59, 0.16), $P = 0.095$; $\log_2$(maternal water As) $\rightarrow$ gestational age $\rightarrow$ birth weight: B (95% CI) = -4.52 (-10.45, 1.28), $P = 0.125$). Furthermore, SEMs allow for the inclusion of cases with missing data on some exogenous variables, and therefore participants with missing data on maternal toenail As concentration could be still be used to estimate parameters for other pathways in the model.

We did not measure gene expression and therefore cannot determine if changes in DNA methylation of *DNMT3A* lead to changes in gene expression. In addition, while it is expected that changes in DNMT3A expression would be reflected in DNA methylation levels of other genes, we did not have epigenome-wide DNA methylation data on this subsample of the birth cohort.

7.7 Conclusion

In this study of a prospective birth cohort of mothers exposed to As though drinking water in Bangladesh, we assessed the relationships between *in utero* As exposure, gestational age, birth weight, and methylation of *DNMT3A*, a gene involved in *de novo* DNA methylation. We show that *in utero* As exposure is negatively associated with gestational age and birth weight. The effect of maternal As exposure on gestational age partially mediated by DNA methylation of *DNMT3A*, and the effect on birth weight is fully mediated by pathways including gestational age. These results provide evidence that *in utero* As exposure affects fetal development through epigenetic dys-

regulation. Further research is need to understand how DNA methylation of *DNMT3A* is associated with gene expression, and in turn if this affects the methylation status and expression of other genes related to infant health outcomes.

7.8 Bibliography

Bailey, K A; Smith, A H; Tokar, E J; Graziano, J H; Kim, K-W; Navasumrit, P; Ruchirawat, M; Thiantanawat, A; Suk, W A, and Fry, R C. Mechanisms underlying latent disease risk associated with early-life arsenic exposure: current research trends and scientific gaps. *Environmental Health Perspectives*, 124(2):170–175, 2016. doi: 10.1289/ehp.1409360.

Bangladesh Bureau of Statistics and United Nation Children's Fund. Bangladesh multiple indicator cluster survey 2012-2013 final report. Technical report, Dhaka, Bangladesh, 2015.

Bjørklund, G; Aaseth, J; Chirumbolo, S; Urbina, M A, and Uddin, R. Effects of arsenic toxicity beyond epigenetic modifications. *Environmental Geochemistry and Health*, 40(3):955–965, 2018. doi: 10.1007/s10653-017-9967-9.

Bloom, M S; Surdu, S; Neamtiu, I A, and Gurzau, E S. Maternal arsenic exposure and birth outcomes: a comprehensive review of the epidemiologic literature focused on drinking water HHS Public Access. *International Journal of Hygiene and Environmental Health*, 217(7):709–719, 2014. doi: 10.1016/j.ijheh.2014.03.004.

Bloom, M S; Neamtiu, I A; Surdu, S; Pop, C; Anastasiu, D; Appleton, A A; Fitzgerald, E F, and Gurzau, E S. Low level arsenic contaminated water consumption and birth outcomes in Romania - an exploratory study. *Reproductive Toxicology*, 59:8–16, 2016. doi: 10.1016/j.reprotox.2015.10.012.

Bohlin, J; Håberg, S E; Magnus, P; Reese, S E; Gjessing, H K; Magnus, M C; Parr, C L; Page, C M; London, S J, and Nystad, W. Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biology*, 17(1):207, 2016. doi: 10.1186/s13059-016-1063-4.

Boyle, E M; Poulsen, G; Field, D J; Kurinczuk, J J; Wolke, D; Alfirevic, Z, and Quigley, M A. Effects of gestational age at birth on health outcomes at 3 and 5 years of age: population based cohort study. *BMJ*, 344:e896, 2012. doi: 10.1136/BMJ.E896.

Brosseau-Liard, P E and Savalei, V. Adjusting incremental fit indices for nonnormality. *Multivariate Behavioral Research*, 49(5):460–470, 2014. doi: 10.1080/00273171.2014.933697.

Brosseau-Liard, Patricia E; Savalei, V, and Li, L. An investigation of the sample performance of two nonnormality corrections for RMSEA. *Multivariate Behavioral Research*, 47(6):904–930, 2012. doi: 10.1080/00273171.2012.715252.

Buncher, C R; Succop, P A, and Dietrich, K N. Structural equation modeling in environmental risk assessment. *Environmental Health Perspectives*, 90:209–13, 1991. doi: 10.1289/ehp.90-1519490.

Concha, G; Vogler, G; Lezcano, D; Nermell, B, and Vahter, M. Exposure to inorganic arsenic metabolites during early human development. *Toxicological Sciences*, 44(2): 185–190, 1998.

Creed, J; Brockhoff, C, and Martin, T. Method 200.8, Revision 5.4: Determination of trace elements in waters and wastes by inductively coupled plasma mass spectrometry. Technical report, Environmental Monitoring Systems Laboratory, Office of Research and Development, US Environmental Protection Agency, Cincinnati, OH, 1994. URL `www.epa.gov`.

Farzan, S F; Karagas, M R, and Chen, Y. *In utero* and early life arsenic exposure in relation to long-term health and disease. *Toxicology and Applied Pharmacology*, 272 (2):384–390, 2013. doi: 10.1016/j.taap.2013.06.030.

Ferm, V H and Hanlon, D P. Constant rate exposure of pregnant hamsters to arsenate during early gestation. *Environmental Research*, 37(2):425–32, 1985.

Freire, C; Amaya, E; Gil, F; Murcia, M; LLop, S; Casas, M; Vrijheid, M; Lertxundi, A; Irizar, A; Fernández-Tardón, G; Castro-Delgado, R V; Olea, N; Fernández, M F, and INMA Project. Placental metal concentrations and birth outcomes: the Environment and Childhood (INMA) project. *International Journal of Hygiene and Environmental Health*, 222(3):468–478, 2019. doi: 10.1016/j.ijheh.2018.12.014.

Gu, T; Lin, X; Cullen, S M; Luo, M; Jeong, M; Estecio, M; Shen, J; Hardikar, S; Sun, D; Su, J; Rux, D; Guzman, A; Lee, M; Qi, L S; Chen, J-J; Kyba, M; Huang, Y; Chen, T; Li, W, and Goodell, M A. DNMT3A and TET1 cooperate to regulate promoter epigenetic landscapes in mouse embryonic stem cells. *Genome Biology*, 19 (1):88, 2018. doi: 10.1186/s13059-018-1464-7.

Gunzler, D; Chen, T; Wu, P, and Zhang, H. Introduction to mediation analysis with structural equation modeling. *Shanghai Archives of Psychiatry*, 25(6):390–4, 2013. doi: 10.3969/j.issn.1002-0829.2013.06.009.

Hall, M; Gamble, M; Slavkovich, V; Liu, X; Levy, D; Cheng, Z; van Geen, A; Yunus, M; Rahman, M; Pilsner, J R, and Graziano, J. Determinants of arsenic metabolism: blood arsenic metabolites, plasma folate, cobalamin, and homocysteine concentrations in maternalnewborn pairs. *Environmental Health Perspectives*, 115(10):1503–1509, 2007. doi: 10.1289/ehp.9906.

Heindel, J J and Vandenberg, L N. Developmental origins of health and disease: a paradigm for understanding disease cause and prevention. *Current Opinion in Pediatrics*, 27(2):248–53, 2015. doi: 10.1097/MOP.0000000000000191.

Hill, D S; Wlodarczyk, B J, and Finnell, R H. Reproductive consequences of oral arsenate exposure during pregnancy in a mouse model. *Birth Defects Research (Part B)*, 83:40–7, 2008. doi: 10.1002/bdrb.20142.

Hood, R D and Bishop, S L. Teratogenic effects of sodium arsenate in mice. *Archives of Environmental Health*, 24(1):62–65, 1972. doi: 10.1080/00039896.1972.10666051.

Hood, Ronald D. Effects of sodium arsenite on fetal development. *Bulletin of Environmental Contamination and Toxicology*, 7(4):216–222, 1972. doi: 10.1007/BF01684401.

Kaneda, M; Okano, M; Hata, K; Sado, T; Tsujimoto, N; Li, E, and Sasaki, H. Essential role for *de novo* DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature*, 429(6994):900–903, 2004. doi: 10.1038/nature02633.

Karagas, M R; Morris, J S; Weiss, J E; Spate, V; Baskett, C, and Greenberg, E R. Toenail samples as an indicator of drinking water arsenic. *Biomarkers and Prevention*, 5:849–852, 1996.

Kile, M L; Houseman, E A; Rodrigues, E; Smith, T J; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Su, L, and Christiani, D C. Toenail arsenic concentrations, GSTT1 gene polymorphisms, and arsenic exposure from drinking water. *Cancer Epidemiology Biomarkers and Prevention*, 14(10):2419–2426, 2005. doi: 10.1158/1055-9965.EPI-05-0306.

Kile, M L; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Cardenas, A; Wright, R O, and Christiani, D C. Effect of prenatal arsenic exposure on DNA methylation and leukocyte subpopulations in cord blood. *Epigenetics*, 9(5):774–82, 2014. doi: 10.4161/epi.28153.

Kile, M L; Cardenas, A; Rodrigues, E; Mazumdar, M; Dobson, C; Golam, M; Quamruzzaman, Q; Rahman, M, and Christiani, D C. Estimating effects of arsenic exposure during pregnancy on perinatal outcomes in a Bangladeshi cohort. *Epidemiology*, 27 (2):1, 2015. doi: 10.1097/EDE.0000000000000416.

Kozul-Horvath, C D; Zandbergen, F; Jackson, B P; Enelow, R I, and Hamilton, J W. Effects of low-dose drinking water arsenic on mouse fetal and postnatal growth and development. *PLoS ONE*, 7(5):e38249, 2012. doi: 10.1371/journal.pone.0038249.

Laine, J E; Bailey, K A; Rubio-Andrade, M; Olshan, A F; Smeester, L; Drobná, Z; Herring, A H; Stýblo, M; García-Vargas, G G, and Fry, R C. Maternal arsenic exposure, arsenic methylation efficiency, and birth outcomes in the Biomarkers of Exposure to ARsenic (BEAR) Pregnancy cohort in Mexico. *Environmental Health Perspectives*, 123(2):186–192, 2015. doi: 10.1289/ehp.1307476.

Maruyama, G. *Basics of Structural Equation Modeling*. SAGE Publications, Thousand Oaks, CA, 1998.

Milton, A H; Hussain, S; Akter, S; Rahman, M; Mouly, T A, and Mitchell, K. A review of the effects of chronic arsenic exposure on adverse pregnancy outcomes. *International Journal of Environmental Research and Public Health*, 14(6), 2017. doi: 10.3390/ijerph14060556.

Monawar Hosain, G M; Chatterjee, N; Begum, A, and Saha, S C. Factors associated with low birthweight in rural Bangladesh. *Journal of Tropical Pediatrics*, 52(2): 87–91, 2005. doi: 10.1093/tropej/fmi066.

Moore, C L; Flanigan, T J; Law, C D; Loukotková, L; Woodling, K A; da Costa, G G; Fitzpatrick, S C, and Ferguson, S A. Developmental neurotoxicity of inorganic arsenic exposure in Sprague-Dawley rats. *Neurotoxicology and Teratology*, 72:49–57, 2019. doi: 10.1016/j.ntt.2019.01.007.

Morrissey, R E and Mottet, N K. Arsenic-induced exencephaly in the mouse and associated lesions occurring during neurulation. *Teratology*, 28(3):399–411, 1983. doi: 10.1002/tera.1420280311.

Nagymajtényi, L; Selypes, A, and Berencsi, G. Chromosomal aberrations and fetotoxic effects of atmospheric arsenic exposure in mice. *Journal of Applied Toxicology*, 5(2): 61–3, 1985.

Non, A L; Binder, A M; Kubzansky, L D, and Michels, K B. Genome-wide DNA methylation in neonates exposed to maternal depression, anxiety, or SSRI medication during pregnancy. *Epigenetics*, 9(7):964–972, 2014. doi: 10.4161/epi.28853.

Oberski, D. lavaan.survey: an R package for complex survey analysis of structural equation models. *Journal of Statistical Software*, 57(1):1–27, 2014. doi: 10.18637/jss.v057.i01. URL `http://www.jstatsoft.org/v57/i01/`.

Okano, M; Bell, D W; Haber, D A, and Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell*, 99(3):247–57, 1999.

Parets, S E; Conneely, K N; Kilaru, V; Fortunato, S J; Syed, T Ali; Saade, G; Smith, A K, and Menon, R. Fetal DNA methylation associates with early spontaneous preterm birth and gestational age. *PloS ONE*, 8(6):e67489, 2013. doi: 10.1371/journal.pone.0067489.

R Core Team. R: A Language and Environment for Statistical Computing. Technical report, R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `https://www.r-project.org/`.

Rahman, A; Granberg, C, and Persson, L-Å. Early life arsenic exposure, infant and child growth, and morbidity: a systematic review. *Archives of Toxicology*, 91(11): 3459–3467, 2017. doi: 10.1007/s00204-017-2061-3.

Rodrigues, E G; Kile, M; Dobson, C; Amarasiriwardena, C; Quamruzzaman, Q; Rahman, M; Golam, M, and Christiani, D C. Maternalinfant biomarkers of prenatal exposure to arsenic and manganese. *Journal of Exposure Science and Environmental Epidemiology*, 25(6):639–648, 2015. doi: 10.1038/jes.2015.45.

Röllin, H B; Channa, K; Olutola, B G, and Odland, J Ø. Evaluation of *in utero* exposure to arsenic in South Africa. *Science of the Total Environment*, 575:338–346, 2017. doi: 10.1016/j.scitotenv.2016.10.044.

Shah, R; Mullany, L C; Darmstadt, G L; Mannan, I; Rahman, S M; Talukder, R Rahman; Applegate, J A; Begum, N; Mitra, D; Arifeen, S E; Baqui, A H, and ProjAHNMo Study Group in Bangladesh. Incidence and risk factors of preterm birth in a rural Bangladeshi cohort. *BMC Pediatrics*, 14(1):112, 2014. doi: 10.1186/1471-2431-14-112.

Sharp, G C; Lawlor, D A; Richmond, R C; Fraser, A; Simpkin, A; Suderman, M; Shihab, H A; Lyttleton, O; McArdle, W; Ring, S M; Gaunt, T R; Davey Smith, G, and Relton, C L. Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: findings from the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, 44(4):1288–1304, 2015. doi: 10.1093/ije/dyv042.

Sun, X; Liu, W; Zhang, B; Shen, X; Hu, C; Chen, X; Jin, S; Jiang, Y; Liu, H; Cao, Z; Xia, W; Xu, S, and Li, Y. Maternal heavy metal exposure, thyroid hormones, and birth outcomes: a prospective cohort study. *Journal of Clinical Endocrinology and Metabolism*, page Epub ahead of print, 2019. doi: 10.1210/jc.2018-02492.

Uysal, F; Ozturk, S, and Akkoyunlu, G. DNMT1, DNMT3A and DNMT3B proteins are differently expressed in mouse oocytes and early embryos. *Journal of Molecular Histology*, 48:417–426, 2017. doi: 10.1007/s10735-017-9739-y.

Vahter, M. Effects of arsenic on maternal and fetal health. *Annual Review of Nutrition*, 29(1):381–399, 2009. doi: 10.1146/annurev-nutr-080508-141102.

Wardlaw, T; Blanc, A; Zupan, J; Ahman, E, and Åhman, E. Low birthweight: country, regional, and global estimates. Technical report, New York, NY, 2004.

Watanabe, D; Suetake, I; Tada, T, and Tajima, S. Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. *Mechanisms of Development*, 118(1-2):187–190, 2002. doi: 10.1016/S0925-4773(02)00242-3.

World Health Organization. Arsenic in drinking-water: background document for the development of WHO guidelines for drinking-water quality. Technical report, Geneva, Switzerland, 2011.

Xu, L; Yokoyama, K; Tian, Y; Piao, F-Y; Kitamura, F; Kida, H, and Wang, P. Decrease in birth weight and gestational age by arsenic among the newborn in Shanghai, China. *[Nihon koshu eisei zasshi] Japanese Journal of Public Health*, 58(2):89–95, 2011.

Zhong, Q; Cui, Y; Wu, H; Niu, Q; Lu, X; Wang, L, and Huang, F. Association of maternal arsenic exposure with birth size: a systematic review and meta-analysis. *Environmental Toxicology and Pharmacology*, 69:129–136, 2019. doi: 10.1016/j.etap.2019.04.007.

7.9 Appendix

Table 7.6: SEM for the mediated effect of postpartum maternal toenail As levels on birth weight through *DNMT3A* DNA methylation.

Pathway	**Effect**	**B**	**(95% CI)**	***P***
DNMT3A				
$\log_2$(Maternal toenail As) → *DNMT3A* DNAm	Direct	0.37	(0.12, 0.62)	0.004
Birth weight				
$\log_2$(Maternal toenail As) → Birth weight	Direct	1.62	(-26.54, 29.78)	0.91
DNMT3A DNAm → Birth weight	Direct	-20.67	(-32.25, -9.10)	<0.001
$\log_2$(Maternal toenail As) → *DNMT3A* DNAm → Birth weight	Indirect	-7.62	(-13.78, -1.46)	0.015

Fit indices calculated from robust estimators: χ^2 P-value = 0.879; root mean square error of approximation (RMSEA) = 0.000; comparative fit index (CFI) = 1.000; Tucker-Lewisnon-normed fit index = 1.036; standardized root mean squared residual = 0.005.
Abbreviation: DNAm. DNA methylation.

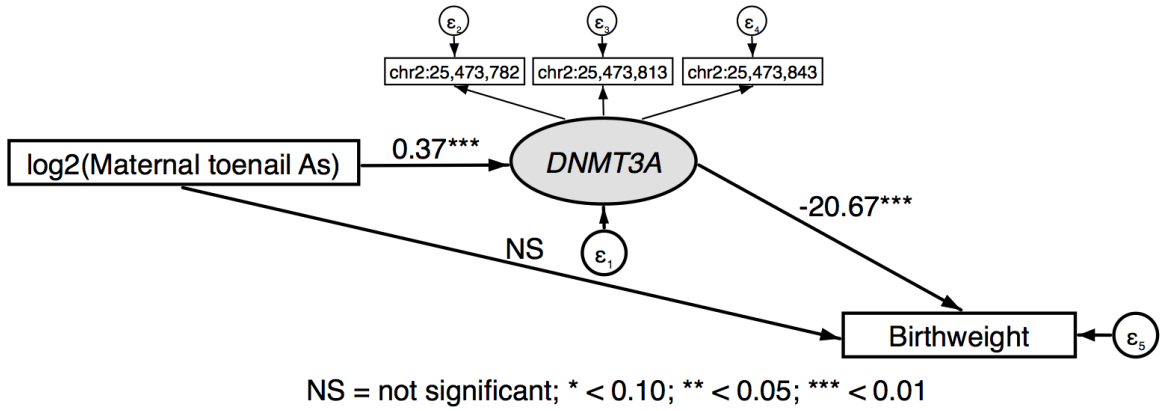


Figure 7.6: SEM for the mediated effect of postpartum maternal toenail As levels on birth weight through *DNMT3A* DNA methylation. Abbreviation: NS, non-significant.

Table 7.7: SEM for the mediated effect of postpartum maternal toenail As levels on gestational age through *DNMT3A* DNA methylation.

Pathway	**Effect**	**B**	**(95% CI)**	***P***
DNMT3A				
$\log_2$(Maternal toenail As) → *DNMT3A* DNAm	Direct	0.39	(0.14, 0.65)	0.002
Gestational age				
$\log_2$(Maternal toenail As) → Gestational age	Direct	-0.25	(-0.43, -0.08)	0.004
DNMT3A DNAm → Gestational age	Direct	-0.12	(-0.18, -0.05)	<0.001
$\log_2$(Maternal toenail As) → *DNMT3A* DNAm → Gestational age	Indirect	-0.05	(-0.08, -0.01)	0.018

Fit indices calculated from robust estimators: χ^2 P-value = 0.122; root mean square error of approximation (RMSEA) = 0.045; comparative fit index (CFI) = 0.998; Tucker-Lewisnon-normed fit index = 0.995; standardized root mean squared residual = 0.010.
Abbreviation: DNAm, DNA methylation.

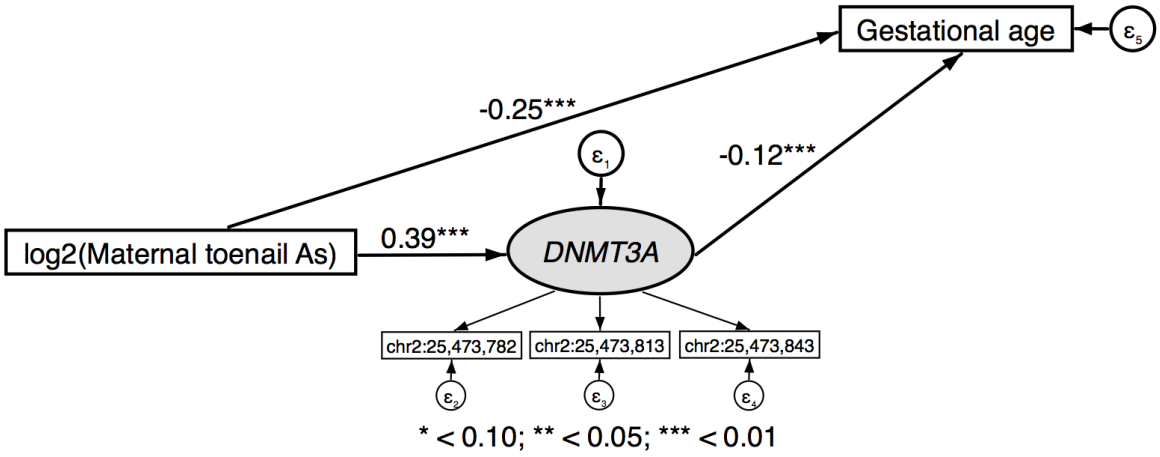


Figure 7.7: SEM for the mediated effect of postpartum maternal toenail As levels on gestational age through *DNMT3A* DNA methylation. Abbreviation: NS, non-significant.

Table 7.8: SEM for the mediated effect of maternal toenail As at enrollment on gestational through *DNMT3A* DNA methylation, and the mediated effect of maternal toenail As at enrollment on birth weight through gestational age and *DNMT3A* DNA methylation, adjusted for infant sex, maternal weight gain, and maternal education.

Pathway	**Effect**	***B***	**(95% CI)**	***P***
DNMT3A				
$\log_2$(Maternal toenail As, enrollment) → *DNMT3A* DNAm	Direct	0.4	(0.14, 0.67)	0.003
Infant sex → *DNMT3A* DNAm	Direct	-0.68	(-1.40, 0.03)	0.063
Maternal weight gain				
$\log_2$(Maternal toenail As, enrollment) → Maternal weight gain	Direct	-0.18	(-0.35, 0.00)	0.046
Gestational age				
$\log_2$(Maternal toenail As, enrollment) → Gestational age	Direct	-0.24	(-0.42, -0.07)	0.007
DNMT3A DNAm → Gestational age	Direct	-0.1	(-0.16, -0.04)	0.001
Maternal weight gain → Gestational age	Direct	0.28	(0.20, 0.36)	<0.001
Maternal education (> primary vs. ≤ primary) → Gestational age	Direct	0.56	(0.16, 0.97)	0.005
$\log_2$(Maternal toenail As, enrollment) → *DNMT3A* DNAm → Gestational age	Indirect	-0.04	(-0.08, -0.01)	0.03
$\log_2$(Maternal toenail As, enrollment) → Maternal weight gain → Gestational age	Indirect	-0.05	(-0.10, 0.00)	0.05
Birth weight				
DNMT3A DNAm → Birth weight	Direct	-11.01	(-21.73, 0.69)	0.053
Gestational age → Birth weight	Direct	77	(58.43, 94.34)	<0.001
Infant sex → Birth weight	Direct	-118.92	(-195.06, -42.45)	0.002
Maternal weight gain → Birth weight	Direct	7.47	(-10.61, 26.20)	0.423
Maternal education → Birth weight	Direct	-60.8	(-142.13, 17.60)	0.135
$\log_2$(Maternal toenail As, enrollment) → *DNMT3A* DNAm → Birth weight	Indirect	-4.44	(-9.76, 0.27)	0.081
$\log_2$(Maternal toenail As, enrollment) → Gestational age → Birth weight	Indirect	-18.25	(-33.66, -5.28)	0.01
$\log_2$(Maternal toenail As, enrollment) → *DNMT3A* DNAm → Gestational age → Birth weight	Indirect	-3.11	(-6.62, -0.70)	0.045
$\log_2$(Maternal toenail As, enrollment) → Maternal weight gain → Birth weight	Indirect	-1.33	(-5.22, 2.29)	0.468
$\log_2$(Maternal toenail As, enrollment) → Maternal weight gain → Gestational age → Birth weight	Indirect	-3.83	(-8.18, 0.01)	0.064

Fit indices: χ^2 *P*-value = 0.045; root mean square error of approximation (RMSEA) = 0.039; comparative fit index (CFI) = 0.994; Tucker-Lewisnon-normed fit index = 0.988; standardized root mean squared residual = 0.032.
Abbreviation: DNAm, DNA methylation.

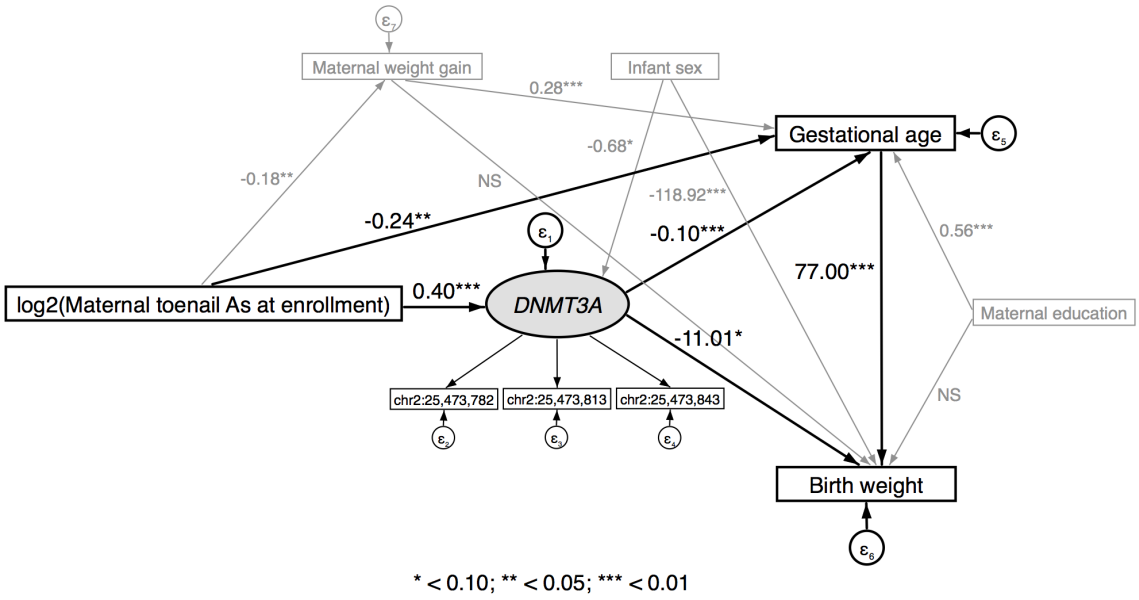


Figure 7.8: SEM for the mediated effect of maternal toenail As at enrollment on gestational age through *DNMT3A* DNA methylation. Abbreviation: NS, non-significant.

Table 7.9: SEM for the mediated effect of maternal water As levels on gestational through *DNMT3A* DNA methylation, and the mediated effect of maternal water As levels on birth weight through gestational age and *DNMT3A* DNA methylation, adjusted for infant sex, maternal weight gain, and maternal education.

Pathway	**Effect**	***B***	**(95% CI)**	***P***
DNMT3A				
$\log_2$(Maternal water As) → *DNMT3A* DNAm	Direct	0.18	(0.06, 0.31)	0.004
Infant sex → *DNMT3A* DNAm	Direct	-0.64	(-1.36, 0.06)	0.076
Maternal weight gain				
$\log_2$(Maternal water As) → Maternal weight gain	Direct	-0.16	(-0.25, -0.07)	0.001
Gestational age				
$\log_2$(Maternal water As) → Gestational age	Direct	-0.06	(-0.13, 0.02)	0.118
DNMT3A DNAm → Gestational age	Direct	-0.11	(-0.17, -0.05)	<0.001
Maternal weight gain → Gestational age	Direct	0.28	(0.20, 0.37)	<0.001
Maternal education (> primary vs. ≤ primary) → Gestational age	Direct	0.55	(0.15, 0.95)	0.007
$\log_2$(Maternal water As) → *DNMT3A* DNAm → Gestational age	Indirect	-0.02	(-0.04, 0.00)	0.031
$\log_2$(Maternal water As) → Maternal weight gain → Gestational age	Indirect	-0.04	(-0.07, -0.02)	0.001
Birth weight				
DNMT3A DNAm → Birth weight	Direct	-11.03	(-21.66, 0.81)	0.054
Gestational age → Birth weight	Direct	77.01	(57.81, 93.99)	<0.001
Infant sex → Birth weight	Direct	-118.87	(-197.12, -41.06)	0.003
Maternal weight gain → Birth weight	Direct	7.46	(-10.73, 25.57)	0.419
Maternal education → Birth weight	Direct	-60.88	(-139.73, 19.15)	0.135
$\log_2$(Maternal water As) → *DNMT3A* DNAm → Birth weight	Indirect	-2.01	(-4.59, 0.16)	0.095
$\log_2$(Maternal water As) → Gestational age → Birth weight	Indirect	-4.52	(-10.45, 1.28)	0.125
$\log_2$(Maternal water As) → *DNMT3A* DNAm → Gestational age → Birth weight	Indirect	-1.53	(-3.28, -0.32)	0.045
$\log_2$(Maternal water As) → Maternal weight gain → Birth weight	Indirect	-1.16	(-3.93, 1.85)	0.419
$\log_2$(Maternal water As) → Maternal weight gain → Gestational age → Birth weight	Indirect	-3.34	(-5.92, -1.30)	0.004

Fit indices: χ^2 P-value = 0.169; root mean square error of approximation (RMSEA) = 0.027; comparative fit index (CFI) = 0.997; Tucker-Lewisnon-normed fit index = 0.994; standardized root mean squared residual = 0.028.
Abbreviation: DNAm, DNA methylation.

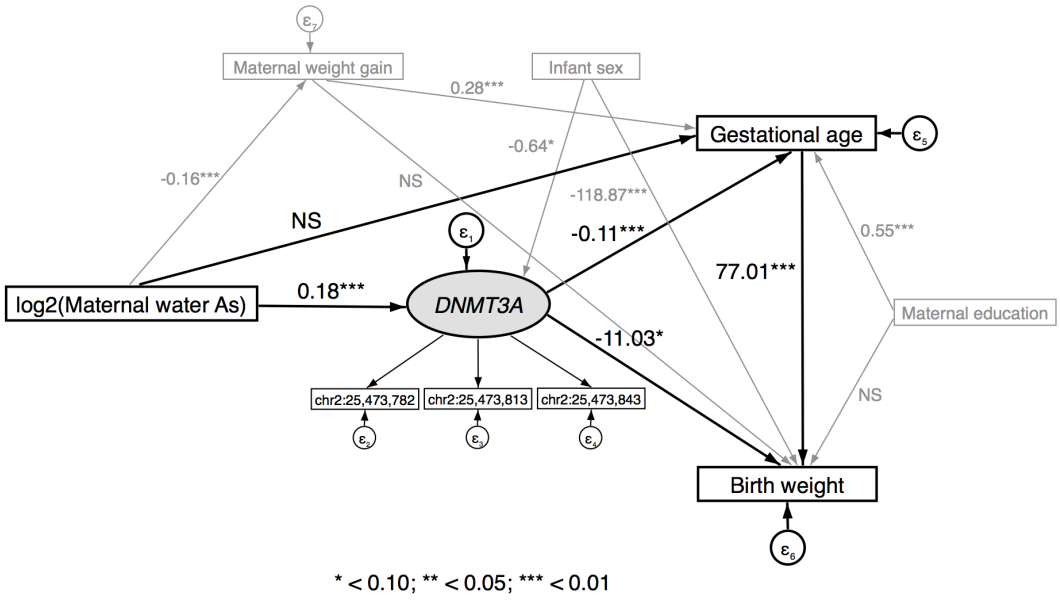


Figure 7.9: SEM for the mediated effect of maternal drinkng water Ason gestational age through *DNMT3A* DNA methylation. Abbreviation: NS, non-significant.

Table 7.10: Linear models to test the associations between DNA methylation of the target CpG site (cg26544247), $\log_2$(postpartum maternal toenail As concentrations), birth weight, and gestational age.

		Crude			Adjusted[a]		
Relationship		***B***	**95% CI**	***P***	***B***	**95% CI**	***P***
$\log_2$(Maternal toenail As) → Gestational age	Exposure → outcome	-0.21	(-0.34, -0.09)	<0.001	-0.16	(-0.27, -0.04)	0.01
$\log_2$(Maternal toenail As) → Birth weight	Exposure → outcome	-4.64	(-32.43, 23.14)	0.743	1.92	(-25.61, 29.45)	0.891
$\log_2$(Maternal toenail As) → cg26544247 DNAm	Exposure → outcome	0.46	(0.16, 0.76)	0.003	0.44	(0.14, 0.74)	0.005
cg26544247 DNAm → Gestational age	Mediator → outcome	-0.1	(-0.15, -0.06)	<0.001	-0.09	(-0.13, -0.04)	<0.001
cg26544247 DNAm → Birth weight	Mediator → outcome	-16.23	(-25.49, -6.96)	<0.001	-12.57	(-24.77, -6.37)	<0.001
Gestational age → Birth weight	Mediator → outcome	79.67	(62.50, 96.84)	<0.001	79.11	(60.74, 97.47)	<0.001

a. Adjusted for infant sex, maternal weight gain, and maternal education. Abbreviation: DNAm, DNA methylation.

Chapter 8

Conclusions and future directions

8.1 Summary of main findings

The overall objective of this dissertation was to investigate nutritional factors that influence arsenic (As) methylation and the relationships between chronic As exposure, epigenetic dysregulation, and adverse health outcomes (Figure 8.1). Specifically, the following aims were addressed: (1) to evaluate overall treatment effects of folic acid (FA) and creatine supplementation on As methylation capacity, and effect modification of treatment effects by baseline status of one-carbon metabolism (OCM) nutrients; (2) examine associations between As exposure and locus-specific DNA methylation in epigenomic-wide association study (EWAS); and (3) to assess mediation of the association between *in utero* As exposure and birth outcomes (i.e., gestational age and birth weight) by DNA methylation of target genes identified in an EWAS as well as the candidate gene DNA methyltransferase 3 alpha (*DNMT3A*). This chapter will summarize the main findings of this dissertation and discuss future research directions.

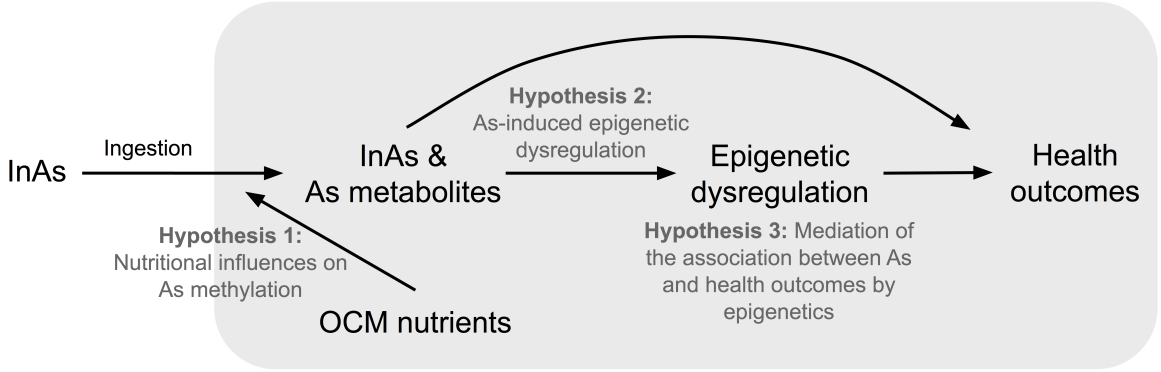


Figure 8.1: Overview of hypotheses.

8.2 Hypothesis 1

Hypothesis 1 Arsenic is methylated using the methyl donor *S*-adenosylmethionine (SAM), which is synthesized by OCM. Folate recruits methyl groups to OCM, and folate

status is positively associated with As methylation capacity (Bozack et al., 2018b). In a previous randomized controlled trial (RCT) of As-exposed adults in Bangladesh with plasma folate < 9 nmol/L, our group found that FA supplementation was associated with a greater increase in the proportion of urinary dimethyl-arsenical species (%DMAs) and decreases in the proportions of urinary inorganic As (%InAs) and monomethyl-arsenical species (%MMAs), and blood As (bAs) and MMAs concentrations compared to placebo (Gamble et al., 2006, 2007). Hypothesis 1 used data from the Folic Acid and Creatine Trial (FACT), a randomized clinical trial of FA (400 or 800 μg/day) and/or creatine supplementation (3 g/day or 3 g creatine and 400 μg FA/day) among As-exposed adults in Bangladesh recruited independent of folate status. Our group previously reported significantly greater decrease in bAs with 800 μg/day FA treatment compared to placebo at week 12 in FACT (Peters et al., 2015a).

Here, we hypothesized that supplementation of FA and creatine would increase As methylation capacity measured by %InAs, %MMAs, %DMAs in urine (*Aim 1a*) and that As methylation capacity would rebound after cessation of FA supplementation (*Aim 1b*). We also hypothesized that baseline nutritional status of OCM-related nutrients, specifically the alternative methyl donor betaine, its precursor choline, folate, and the cofactor vitamin B_{12}, would modify FA and creatine treatment effects on changes in total homocysteine and guanidinoacetate (GAA) concentrations (biomarkers of OCM and endogenous creatine synthesis, respectively), total bAs concentrations, and urinary As metabolite proportions and methylation indices.

Chapter 3 reported overall treatment effects of FA and creatine supplementation on changes in urinary As metabolite proportions. We observed significantly greater mean within-person decreases %InAs and %MMAs and increase in %DMAs among all groups receiving FA supplementation at weeks 6 and 12 compared to placebo ($P < 0.05$). The decrease in %MMAs, however, was greater in the 800 μg/day FA

group than the 400 μg/day FA ($P < 0.05$), suggesting that the observed decrease in bAs was driven larger increase in As methylation capacity in the 800 μg/day FA group, leading to an increase in urinary As excretion. When participants were stratified by median choline and betaine concentrations at baseline (reported in **Chapter 4**), we observed a trend towards greater FA treatment effects among participants with levels below the median of both nutrients compared to participants above the median. The mean within-person increase in urinary %DMAs with 400 and 800 μg/day compared to placebo was significantly different between the high and low betaine strata after correcting for multiple tests ($FDR < 0.05$). From weeks 12 to week 24, half of the participants in groups receiving FA were switched to placebo. At week 24, %InAs and %DMAs were not significantly different than baseline levels among participants who discontinued FA supplementation, suggesting a rebound in As methylation capacity with cessation of FA supplementation.

Endogenous creatine biosynthesis is a major consumer of SAM (Brosnan et al., 2011) and urinary creatinine, a biomarker of dietary creatine and product of creatine biosynthesis, is positively associated with As methylation capacity (Basu et al., 2011; Gamble et al., 2005b, 2006; Hall et al., 2007, 2009; Kile et al., 2009). In FACT, greater decreases in GAA with creatine supplementation compared to placebo indicated downregulation of creatine biosynthesis (Peters et al., 2015b). As described in **Chapter 3**, we observed a significantly greater mean within-person decreases in %MMAs with creatine supplementation compared to placebo at weeks 1, 6, and 12; however, mean within-person changes in %InAs and %DMAs did not differ significantly between the creatine and placebo groups. We also investigated creatine treatment effects stratified by median values of OCM-related nutrients (**Chapter 4**). The mean within-person decrease in urinary %MMAs at week 12 was significantly greater than placebo among participants with baseline choline concentrations below the median, but did not dif-

fer from placebo among participants with choline concentrations above the median. This treatment effect was different between strata at $P < 0.10$. Choline can be obtained through diet or synthesized from phosphatidylcholine (PC), which in turn can by biosynthesized using three molecules of SAM for each PC molecule. In addition to choline's role as a precursor to the methyl donor betaine, choline is involved in range of biological functions (Zeisel, 2009). Considering the biological roles of choline in addition to acting as a precursor to betaine, we hypothesize that low choline status may deplete due SAM due to increased PC biosynthesis, therefore resulting in a low-SAM scenario where creatine treatment effects on %MMAs are more pronounced.

Overall, results from FACT reported in **Chapters 3 and 4** provide evidence of the associations between OCM-related nutrients and As methylation capacity. Specifically, FA and creatine supplementation may increase As methylation capacity by increasing the availability of SAM, and treatment effects may be greater among individuals with low betaine and choline status, respectively.

8.3 Hypothesis 2

Adverse health outcomes associated with chronic As exposure, including cancer mortality (Roh et al., 2018; Smith et al., 2018) and lung disease (Steinmaus et al., 2016) persists decades after exposure has been reduced. This suggests that epigenetic dysregulation, including changes in DNA methylation, may be one biological mechanism linking As exposure and health outcomes (Bailey et al., 2016). In epigenome-wide association studies (EWAS) using blood DNA, DNA methylation at individual loci has been associated with As exposure *in utero* (Broberg et al., 2014; Kaushal et al., 2017; Kile et al., 2014; Rojas et al., 2015) and in adulthood (Ameer et al., 2017; Argos et al., 2015; Demanelis et al., 2019).

We conducted an EWAS to investigate the association between urinary As

levels and DNA methylation measured using the in the Illumina MethylationEPIC (850K) array in blood samples from the Strong Heart Study (SHS), a population-based prospective cohort of American Indian adults with low-to-moderate levels of As exposure, with the hypothesis that As exposure is associated with DNA methylation at individual loci and regions (*Aim 2a*). Within this EWAS, we also conducted gene ontology (GO) enrichment analysis to identify biological and molecular functions over-represented among differentially methylated positions (DMPs) (*Aim 2b*). Results of Hypothesis 2 were described in **Chapter 5**.

We identified 20 DMPs significantly associated with urinary As levels at $FDR < 0.05$; five DMPs remained significant after adjusting for multiple tests using a Bonferroni correction. Overall, there was a trend toward a positive association between urinary As levels and DNA methylation; As levels were positively associated with methylation at18 of the 20 FDR-significant CpGs, and As was positively associated with methylation at 58% of all probes. The top significant CpG, cg06690548, was located in solute carrier family 7 member 11 (*SLC7A11*), part of the amino-acid transporter cystine:glutamate antiporter system x_c^-, which is involved in biosynthesis of the endogenous antioxidant glutathione (GSH) (Conrad and Sato, 2012; Lim and Donaldson, 2011). Nominally significant results from GO enrichment analysis included related terms associated with the transport of cysteine, confirming the involvement of SLC7A11A in these pathways related to GSH biosynthesis. Arsenic exposure has previously been associated with blood GSH concentration in a cross-section study (Hall et al., 2013). Arsenic may deplete GSH binding to GSH and/or through the induction of reactive oxygen species; however, our results also suggest that As may affect GSH biosynthesis through an epigenetic mechanism. Additional FDR-significant CpGs were located in biologically relevant genes including leucine rich repeat and Ig domain containing 3 (*LINGO3*; cg22294740 and cg08059112), a transmembrane protein that

has been identified as a gene hub in metastatic melanoma tumors (Wang et al., 2018); casein kinase 1 delta (*CSNK1D*; cg20493718 and cg21369801), which may be involved in cellular processes including control of DNA replication and repair, apoptosis, and circadian rhythm, and regulation of the tumor suppressor protein p53 (Schittek and Sinnberg, 2014); and hexokinase 1 (*HK1*; cg14595618), a protein coding gene involved in glucose metabolism. In sensitivity analyses stratified by diabetes status, cg14595618 was not associated with As levels at a nominal $P < 0.05$. In a separate linear model, we also found significant effect modification of the association between As levels and cg14595618 methylation by diabetes status.

We also identified one FDR-significant differentially methylated region (DMR; chr11:2,322,050-2,323,247) including the open reading frame *C11orf21* and tetraspanin 32 (*TSPAN32*). Genetic alterations in C11orf21/TSPAN32 have been associated with Beckwith-Wiedemann syndrome, a condition characterized by abnormal growth and tumors in childhood (Koufos et al., 1989) and with chronic lymphocytic leukemia (Berndt et al., 2013).

We also assessed significance of DMPs previously identified in EWAS. Although no consistent epigenetic signature of As exposure has been identified, among 145 DMPs identified in studies including adults, 12 achieved nominal significance in our study.

In summary, **Chapter 5** reports results from an EWAS of the association between chronic As exposure and DNA methylation. To our knowledge, this is the largest EWAS of As exposure to date, and one of the first using the 850K microarray. We identified 20 FDR-significant DMPs, including loci annotated to genes biologically relevant to As exposure.

8.4 Hypothesis 3

In utero As exposure may have teratogenic effects. Overall, epidemiological studies have reported a trend toward decreased birth weight associated with higher *in utero* As exposures (Bloom et al., 2014; Milton et al., 2017; Zhong et al., 2019); however, fewer studies have investigated associations with gestational age. Observed reductions in birth weight may be due to shortened gestational period and/or intrauterine growth restriction, factors which are associated with different health impacts in infancy and later in life (Wardlaw et al., 2004). Furthermore, environmental exposures during embryogenesis, a time of cellular differentiation and epigenetic reprogramming, may result in epigenetic dysregulation and increased disease risk (Heindel and Vandenberg, 2015), suggesting that there is an epigenetic link between *in utero* As exposure and birth outcomes. Therefore, understanding the etiology of these birth outcome has public health implications.

In a birth cohort in Bangladesh, it has previously been shown that the inverse association between *in utero* As exposure and birth weight is fully mediated by decreases in gestational age and maternal weight gain (Kile et al., 2015). **Chapters 6 and 7** report further analyses from this cohort. We hypothesized that the associations between *in utero* As exposure and birth outcomes, specifically birth weight and gestational age, are mediated by changes in DNA methylation measured in cord blood. In an EWAS, we identified CpGs associated with maternal water As levels and gestational age. Among top CpGs, three were selected for bisulfite pyrosequencing in a separate subset of cohort participants to validate mediation of the association between *in utero* As exposure and gestational age by DNA methylation (*Aim 3a*). We also assessed mediation of the associations between *in utero* As exposure, birth weight, and gestational age by DNA methylation of DNA methyltransferase 3 alpha (*DNMT3A*), which codes for an enzyme responsible for *de novo* DNA methylation and a prime candidate gene as a mediator

between prenatal exposures and birth outcomes (*Aim 3b*).

In **Chapter 6** we introduce an experimental approach for assessing mediation of the association between *in utero* exposures and birth outcomes by DNA methylation including a discovery phase using an epigenome-wide approach and a validation phase using a target CpG approach. In the discovery phase (N = 44), the association between maternal water As levels and gestational age was fully mediated by DNA methylation of the top 10 CpGs associated with both variables. In a validation phase, DNA methylation of three top CpGs located in *miR124-3*, *GNAL*, and *MCC* was measured in cord blood of 569 infants. In structural equation models (SEMs), methylation levels of each target CpG and neighboring methylated loci were used to construct latent variables representing DNA methylation of each gene. In individual SEMs, there were significant indirect effects of maternal water As levels on gestational age through DNA methylation of *miR124-3* and *MCC*; the indirect effect through DNA methylation of *GNAL* was not significant. We constructed an adjusted SEM including latent variables representing DNA methylation of *miR124-3* and *MCC*. Mediation of the association between *in utero* As exposure and gestational age by DNA methylation of *miR124-3* was borderline significant ($P = 0.06$); DNA methylation of *MCC* did not act as a mediator.

In **Chapter 7** we reported results of an analysis of DNA methylation of the target gene *DNMT3A* as a mediator of the association between *in utero* As exposure and birth outcomes. In an adjusted SEM including birth weight and gestational age, maternal toenail As levels were associated with *DNMT3A* methylation and gestational age, and *DNMT3A* methylation was associated with both birth weight and gestational age. There was a significant indirect effect of As on gestational age through *DNMT3A* methylation, the indirect effect on birth weight was borderline significant ($P = 0.082$). However, the indirect effects of maternal toenail As levels on birth weight through all

pathways including gestational age were statistically significant. A doubling in maternal toenail As concentration had a total effect of a decrease in gestational age of 2.1 days and a decrease in birth weight of 28.9 g.

Results from **Chapters 6 and 7** provide evidence that *in utero* As exposure is associated with adverse birth outcomes and with DNA methylation in cord blood, and that alterations in DNA methylation act as a mediator between As exposure and birth outcomes. In summary, we found that *in utero* As exposure is associated with shortened gestational age, and this association is mediated by changes in DNA methylation of *miR124-3*, a gene identified using an EWAS approach, and *DNMT3*, a selected candidate gene. *In utero* As exposure was not directly associated with birth weight, but in analyses of *DNMT3A* methylation, As exposure had an indirect effect on birth weight through pathways including gestational age. These findings support the developmental origins of health and disease hypothesis that environmental exposures during embryogenesis, a time of epigenetic reprogramming, may result in epigenetic dysregulation and increased disease risk (Heindel and Vandenberg, 2015).

8.5 Future directions

Chapters 3 and 4 (Hypothesis 1), provided evidence from FACT, an RCT, to support observational studies reporting positive associations between folate status and As methylation capacity. We also found that FA supplementation may have more profound effects on As methylation capacity among individual with low status of OCM-related nutrients, particularly the alternate methyl donor betaine. Although mitigation of As exposure is the primary means to prevent As-related health outcomes, our results suggest that nutritional interventions may reduce the adverse effects of chronic As exposure, particularly in areas where nutrient deficiencies are common. FA fortification has been shown to decrease the prevalence of folate deficiency (Barnabé et al., 2015;

Odewole et al., 2013) and has been mandated in 87 countries (Zimmerman and Lu, 2015). However, FA fortification has not been implemented in Bangladesh, where the prevalence of folate deficiency is common (Gamble et al., 2005a). Further research is needed to assess the feasibility of nutritional interventions to increase As methylation capacity in countries with high rates of As exposure and nutrient deficiencies such as Bangladesh.

Further research is needed to understand the association between creatine and As methylation capacity. Although urinary creatinine has been consistently associated with As methylation capacity (i.e., lower %InAs and higher %DMAs in urine) in cross-sectional analyses (Basu et al., 2011; Bozack et al., 2018a; Gamble et al., 2005b, 2006; Hall et al., 2007, 2009; Kile et al., 2009; Pilsner et al., 2009), we observed limited treatment effects of creatine supplementation; creatine supplementation was associated with a decrease in %MMAs in urine at weeks 1, 6, and 12, but was not associated with changes in %InAs or %DMAs. Stratified analyses indicated that this treatment effect was driven by changes in %MMAs among participants with low choline status. There are several possible explanations for these results. Low choline may deplete SAM by upregulating PEMT; creatine may have greater treatment effects in the case of low SAM concentrations. Alternatively, creatine may primarily affect OCM in the liver and treatment effects may be attenuated over time due to long-range allosteric regulation of hepatic SAM concentrations, a scenario supported by the observed plateau in creatine treatment effects. Observed cross-sectional associations between urinary creatinine and As metabolite proportion may be due in part to renal tube reabsorption of InAs when urine is more concentrated (Ginsburg and Lotspeich, 1963). To fully understand the relationship between creatine and As methylation capacity, additional research is needed to determine the effects of creatine supplementation on liver-specific SAM concentrations as well as the potential for reabsorption of As metabolites from

tubular fluid.

Our study did not examine long-term effects of increased As methylation capacity on health outcomes. A higher proportion of %MMAs in urine has been consistently associated with higher risk of adverse health outcomes (Ahsan et al., 2007; Chen et al., 2013, 2003; Gilbert-Diamond et al., 2013; Hsueh et al., 1997; Huang et al., 2008, 2006; Laine et al., 2015; Lindberg et al., 2008; López-Carrillo et al., 2014; Melak et al., 2014; Pu et al., 2007; Steinmaus et al., 2006, 2010; Tseng et al., 2005; Wu et al., 2006; Yu et al., 2000), suggesting that increasing As methylation capacity through nutritional interventions would decrease the risk of adverse health outcomes. However, epidemiological studies have not yet established how changes in As methylation capacity may affect risks later in life, particularly for health outcomes with long latency periods.

Chapter 5 (Hypothesis 2) presented results of an EWAS of the association between urinary As and DNA methylation among adults with low-moderate As exposure. Specifically, we assessed associations between total urinary As levels and DNA methylation of autosomes. Within this study, further analysis may investigate associations between As exposure and DNA methylation of sex chromosomes, as well as associations between As exposure and differential variability of DNA methylation and associations between As metabolites and epigenome-wide DNA methylation.

In the overall study population, we observed a significant association between As exposure and DNA methylation of cg14595618, located in *HK1*, a protein coding gene involved in glucose metabolism. However, in sensitivity analyses, cg14595618 methylation was not significantly associated with As exposure among participants without diabetes, and we found a significant interaction between As levels and diabetes status on cg14595618 methylation. Although we removed probes associated with SNPs with minor allele frequencies < 0.05, these results may be driven by genetic variation (e.g., SNP frequency may vary in this study population, or unidentified SNPs may be

present), resulting in methylation quantitative trait loci. In previous analysis of this study population, urinary As concentration was not associated with diabetes incidence; however, %MMAs measured in urine was associated with a lower risk of diabetes (Kuo et al., 2015). Further research is needed to understand genetic factors and DNA methylation pattern in this genomic region, as well the relationship between As exposure, As methylation capacity, and glucose metabolism.

Overall, previous EWAS have found that As exposure *in utero* and in adulthood is associated with locus-specific DNA methylation (Argos, 2015). However, top signals differ among studies, and a specific epigenetic signature of As exposure has not been established. Variation in EWAS results may be due to a variety of methodological differences including population studied, age and duration of As exposure, measure and level of exposure, tissue type, and platform used to quantify DNA methylation. Variation may also be due in part to differences in analytical methods, including data quality control and processing and statistical tests performed. Further efforts are needed to understand degree of overlap between differentially methylated positions (DMPs) identified by previous EWAS, including comparison of all nominally significant DMPs across studies. Future EWAS efforts should also seek to apply standard statistical methods to increase comparability across studies.

Chapters 6 and 7 reported results from mediation analyses of the association between *in utero* As exposure and birth outcomes by DNA methylation measured in cord blood. We found significant mediation by DNA methylation of CpGs identified through a discovery phase using an EWAS approach, as well as CpGs located in the candidate gene DNA methyltransferase 3 alpha (*DNMT3A*). While these results support the hypothesis of fetal origins of health and disease later in life, further research is needed to asses if observed changes in DNA methylation persist and act as mediators with health outcomes later in life, particularly health outcomes associated with preterm

birth and low birth weight.

In **Chapter 7**, we found a significant indirect effect of *in utero* As exposure on gestational age through DNA methylation of *DNMT3A*; the indirect effect on birth weight was borderline significant. DNMT3A is involved in establishing *de novo* DNA methylation during embryogenesis, a time epigenetic reprogramming. DNA methylation of *DNMT3A* is expected to affect expression of this gene, which may in turn affect DNA methylation and expression of other genes and subsequently health outcomes. However, further research is needed to understand these potential biological mechanisms. Specifically, under the assumption that DNA methylation of *DNMT3A* is associated with levels of gene expression, it is not known if this impacts DNA methylation at the global and/or locus-specific levels.

Chapters 5, 6, and 7 examined relationships between As exposure and epigenetic dysregulation. In Chapters 6 and 7, we were also able to assess mediation of the association between As exposure and health outcomes. However, these studies were limited by lack of data on gene expression. Gene expression data can also help understand the potential of DNA methylation to act as a biological mediator of the association between As exposure and health outcomes.

8.6 Conclusion

The overall objectives of this dissertation were to (1) understand factors that may influence susceptibility to As-induced health outcomes, specifically nutritional influences of As methylation and the potential of nutritional interventions to increase As methylation capacity, and to understand (2) As-induced epigenetic dysregulation and (3) its mediating role in the association between As exposure and health outcomes.

In an RCT, we demonstrated that supplementation of FA and creatine, two nutrients involved in OCM, increase As methylation capacity, and that supplementation

may have more profound treatment effects among individuals with low levels of other OCM-related nutrients. Further research is needed to fully understand the relationships between various OCM nutrients and As methylation. However, our results have important implications for potential nutritional interventions to reduce As-induced health outcomes, particularly in countries such as Bangladesh with high prevalence of chronic As exposure and nutritional deficiencies and without FA fortification.

This dissertation also investigated associations between As exposure and epigenetic dysregulation in diverse populations. In the largest EWAS of As exposure to date, we identified novel differentially methylated loci associated with As exposure among a cohort of adults in the US with low-moderate levels of exposure. In addition, using data from a prospective birth cohort in Bangladesh, we found significant mediation of the association between *in utero* As exposure and birth outcomes by DNA methylation. Taken together, these results provide evidence that chronic As exposure is associated with epigenetic dysregulation, and that changes in the epigenome may mediate the association between As exposure and adverse health effects. However, further research is needed to understand the role of additional factors, including genetic variability and altered gene expression, in the pathway from As exposure to health outcomes.

The results presented in this dissertation contribute the body of research supporting the hypotheses that individual factors, including nutritional status, influence As methylation capacity, and that a mode of action of As toxicity may include epigenetic dysregulation. Although further research is needed to fully understand the biological mechanisms underlying As toxicity, these finding may help inform public health interventions designed for specific populations to reduce the adverse health effects of chronic As exposure.

8.7 Bibliography

Ahsan, H; Chen, Y; Kibriya, M G; Slavkovich, V; Parvez, F; Jasmine, F; Gamble, M V, and Graziano, J H. Arsenic metabolism, genetic susceptibility, and risk of premalignant skin lesions in Bangladesh. *Cancer Epidemiology Biomarkers and Prevention*, 16(6):1270–1278, 2007. doi: 10.1158/1055-9965.EPI-06-0676.

Ameer, S S; Engström, K; Hossain, M B; Concha, G; Vahter, M, and Broberg, M. Arsenic exposure from drinking water is associated with decreased gene expression and increased DNA methylation in peripheral blood. *Toxicology and Applied Pharmacology*, 321:57–66, 2017. doi: 10.1016/J.TAAP.2017.02.019.

Argos, M. Arsenic exposure and epigenetic alterations: recent findings based on the Illumina 450K DNA methylation array. *Current Environmental Health Reports*, 2(2): 137–44, 2015. doi: 10.1007/s40572-015-0052-1.

Argos, M; Chen, L; Jasmine, F; Tong, L; Pierce, B L; Roy, S; Paul-Brutus, R; Gamble, M V; Harper, K N; Parvez, F; Rahman, M; Rakibuz-Zaman, M; Slavkovich, V; Baron, J A; Graziano, J H; Kibriya, M G, and Ahsan, H. Gene-specific differential DNA methylation and chronic arsenic exposure in an epigenome-wide association study of adults in Bangladesh. *Environmental Health Perspectives*, 123(1):64–71, 2015. doi: 10.1289/ehp.1307884.

Bailey, K A; Smith, A H; Tokar, E J; Graziano, J H; Kim, K-W; Navasumrit, P; Ruchirawat, M; Thiantanawat, A; Suk, W A, and Fry, R C. Mechanisms underlying latent disease risk associated with early-life arsenic exposure: current research trends and scientific gaps. *Environmental Health Perspectives*, 124(2):170–175, 2016. doi: 10.1289/ehp.1409360.

Barnabé, A; Aléssio, A; Bittar, L; de Moraes Mazetto, B; Bicudo, A M; de Paula, E V; Höehr, N, and M, Annichino-Bizzacchi J. Folate, vitamin B12 and homocysteine status in the post-folic acid fortification era in different subgroups of the Brazilian population attended to at a public health care center. *Nutrition Journal*, 14(1):19, 2015. doi: 10.1186/s12937-015-0006-3.

Basu, A; Mitra, S; Chung, J; Guha Mazumder, DN; Ghosh, N; Kalman, D; von Ehrenstein, O S; Steinmaus, C; Liaw, J, and Smith, A H. Creatinine, diet, micronutrients, and arsenic methylation in West Bengal, India. *Environmental Health Perspectives*, 119(9):1308–1313, 2011. doi: 10.1289/ehp.1003393.

Berndt, S I; Skibola, C F; Joseph, V; Camp, N J; Nieters, A; Wang, Z; Cozen, W; Monnereau, A; Wang, S S, and Kelly, R S. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nature Genetics*, 45(8):868–876, 2013. doi: 10.1038/ng.2652.

Bloom, M S; Surdu, S; Neamtiu, I A, and Gurzau, E S. Maternal arsenic exposure and birth outcomes: a comprehensive review of the epidemiologic literature focused on

drinking water HHS Public Access. *International Journal of Hygiene and Environmental Health*, 217(7):709–719, 2014. doi: 10.1016/j.ijheh.2014.03.004.

Bozack, A K; Hall, M N; Liu, X; Ilievski, V; Lomax-Luu, A M; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T, and Graziano, J H. Folic acid supplementation enhances arsenic methylation: results from a folic acid and creatine supplementation randomized controlled trial in Bangladesh. *American Journal of Clinical Nutrition*, 109(2):380–391, 2018a. doi: 10.1093/ajcn/nqy148.

Bozack, A K; Saxena, R, and Gamble, M V. Nutritional influences on one-carbon metabolism: effects on arsenic methylation and toxicity. *Annual Review of Nutrition*, 38:401–429, 2018b. doi: 10.1146/annurev-nutr-082117-051757.

Broberg, K; Ahmed, S; Engström, K; Hossain, M B; Jurkovic Mlakar, S; Bottai, M; Grandér, M; Raqib, R, and Vahter, M. Arsenic exposure in early pregnancy alters genome-wide DNA methylation in cord blood, particularly in boys. *Journal of Developmental Origins of Health and Disease*, 5(4):288–98, 2014. doi: 10.1017/S2040174414000221.

Brosnan, J T; da Silva, R P, and Brosnan, M E. The metabolic burden of creatine synthesis. *Amino Acids*, 40(5):1325–31, 2011. doi: 10.1007/s00726-011-0853-y.

Chen, Y-A; Lemire, M; Choufani, S; Butcher, D T; Grafodatskaya, D; Zanke, B W; Gallinger, S; Hudson, T J, and Weksberg, R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*, 8(2):203–9, 2013. doi: 10.4161/epi.23470.

Chen, Y-C; Su, H-J J; Guo, Y-L L; Hsueh, Y-M; Smith, T J; Ryan, L M; Lee, M-S, and Christiani, D C. Arsenic methylation and bladder cancer risk in Taiwan. *Cancer Causes and Control*, 14(4):303–10, 2003. doi: 10.1023/A:1023905900171.

Conrad, M and Sato, H. The oxidative stress-inducible cystine/glutamate antiporter, system x_c^-: cystine supplier and beyond. *Amino Acids*, 42(1):231–246, 2012. doi: 10.1007/s00726-011-0867-5.

Demanelis, K; Argos, M; Tong, L; Shinkle, J; Sabarinathan, M; Rakibuz-Zaman, M; Sarwar, G; Shahriar, H; Islam, T; Rahman, M; Yunus, M; Graziano, J H.; Broberg, K; Engström, K; Jasmine, F; Ahsan, H, and Pierce, B L. Association of arsenic exposure with whole blood DNA methylation: an epigenome-wide study of Bangladeshi adults. *Environmental Health Perspectives*, 127(5), 2019. doi: 10.1289/EHP3849.

Gamble, M V; Ahsan, H; Liu, X; Factor-Litvak, P; Ilievski, V; Slavkovich, VV; Parvez, F, and Graziano, J H. Folate and cobalamin deficiencies and hyperhomocysteinemia in Bangladesh. *American Journal of Clinical Nutrition*, 81(6):1372–1377, 2005a.

Gamble, M V; Liu, X; Ahsan, H; Pilsner, R; Ilievski, V; Slavkovich, V; Parvez, F; Levy, D; Factor-Litvak, P, and Graziano, J H. Folate, homocysteine, and arsenic metabolism in arsenic-exposed individuals in Bangladesh. *Environmental Health Perspectives*, 113(12):1683–8, 2005b. doi: 10.1289/ehp.8084.

Gamble, M V; Liu, X; Ahsan, H; Pilsner, J R; Ilievski, V; Slavkovich, V; Parvez, F; Chen, Y; Levy, D; Factor-Litvak, P, and Graziano, J H. Folate and arsenic metabolism: a double-blind, placebo-controlled folic acidsupplementation trial in Bangladesh. *American Journal of Clinical Nutrition*, 84(5):1093–101, 2006. doi: 10.1093/ajcn/84.5.1093.

Gamble, M V; Liu, X; Slavkovich, V; Pilsner, J R; Ilievski, V; Factor-Litvak, P; Levy, D; Alam, S; Islam, M; Parvez, F; Ahsan, H, and Graziano, J H. Folic acid supplementation lowers blood arsenic. *American Journal of Clinical Nutrition*, 86(4):1202–9, 2007. doi: 10.1093/ajcn/86.4.1202.

Gilbert-Diamond, D; Li, Z; Perry, A E; Spencer, S K; Gandolfi, A J, and Karagas, M R. A population-based casecontrol study of urinary arsenic species and squamous cell carcinoma in New Hampshire, USA. *Environmental Health Perspectives*, 121(10): 1154–1160, 2013. doi: 10.1289/ehp.1206178.

Ginsburg, J M and Lotspeich, W D. Interrelations of arsenate and phosphate transport in the dog kidney. *American Journal of Physiology*, 205(4):707–714, 1963. doi: 10.1152/ajplegacy.1963.205.4.707.

Hall, M; Gamble, M; Slavkovich, V; Liu, X; Levy, D; Cheng, Z; van Geen, A; Yunus, M; Rahman, M; Pilsner, J R, and Graziano, J. Determinants of arsenic metabolism: blood arsenic metabolites, plasma folate, cobalamin, and homocysteine concentrations in maternalnewborn pairs. *Environmental Health Perspectives*, 115(10):1503–1509, 2007. doi: 10.1289/ehp.9906.

Hall, M N; Liu, X; Slavkovich, V; Ilievski, V; Pilsner, J R; Alam, S; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate, cobalamin, cysteine, homocysteine, and arsenic metabolism among children in Bangladesh. *Environmental Health Perspectives*, 117(5), 2009. doi: 10.1289/ehp.0800164.

Hall, M N; Niedzwiecki, M; Liu, X; Harper, K N; Alam, S; Slavkovich, V; Ilievski, V; Levy, D; Siddique, A B; Parvez, F; Mey, J L; van Geen, A; Graziano, J, and Gamble, M V. Chronic arsenic exposure and blood glutathione and glutathione disulfide concentrations in Bangladeshi adults. *Environmental Health Perspectives*, 121(9): 1068–74, 2013. doi: 10.1289/ehp.1205727.

Heindel, J J and Vandenberg, L N. Developmental origins of health and disease: a paradigm for understanding disease cause and prevention. *Current Opinion in Pediatrics*, 27(2):248–53, 2015. doi: 10.1097/MOP.0000000000000191.

Hsueh, Y M; Chiou, H Y; Huang, Y L; Wu, W L; Huang, C C; Yang, M H; Lue, L C; Chen, G S, and Chen, C J. Serum beta-carotene level, arsenic methylation capability, and incidence of skin cancer. *Cancer Epidemiology and Prevention Biomarkers*, 6(8): 589–96, 1997.

Huang, Y-K; Tseng, C-H; Huang, Y-L; Yang, M-H; Chen, C-J, and Hsueh, Y-M. Arsenic methylation capability and hypertension risk in subjects living in arseniasis-hyperendemic areas in southwestern Taiwan. *Toxicology and Applied Pharmacology*, 218:135–142, 2006. doi: 10.1016/j.taap.2006.10.022.

Huang, Y-K; Huang, Y-L; Hsueh, Y-M; Yang, M-H; Wu, M-M; Chen, S-Y; Hsu, L-I, and Chen, C-J. Arsenic exposure, urinary arsenic speciation, and the incidence of urothelial carcinoma: a twelve-year follow-up study. *Cancer Causes and Control*, 19 (8):829–839, 2008. doi: 10.1007/s10552-008-9146-5.

Kaushal, A; Zhang, H; Karmaus, W J J; Everson, T M; Marsit, C J; Karagas, M R; Tsai, S-F; Wen, H-J, and Wang, S-L. Genome-wide DNA methylation at birth in relation to *in utero* arsenic exposure and the associated health in later life. *Environmental Health*, 16(1):50, 2017. doi: 10.1186/s12940-017-0262-0.

Kile, M L; Hoffman, E; Hsueh, Y-M; Afroz, S; Quamruzzaman, Q; Rahman, M; Mahiuddin, G; Ryan, L, and Christiani, D C. Variability in biomarkers of arsenic exposure and metabolism in adults over time. *Environmental Health Perspectives*, 117(3):455–460, 2009. doi: 10.1289/ehp.11251.

Kile, M L; Houseman, E A; Baccarelli, A A; Quamruzzaman, Q; Rahman, M; Mostofa, G; Cardenas, A; Wright, R O, and Christiani, D C. Effect of prenatal arsenic exposure on DNA methylation and leukocyte subpopulations in cord blood. *Epigenetics*, 9(5):774–82, 2014. doi: 10.4161/epi.28153.

Kile, M L; Cardenas, A; Rodrigues, E; Mazumdar, M; Dobson, C; Golam, M; Quamruzzaman, Q; Rahman, M, and Christiani, D C. Estimating effects of arsenic exposure during pregnancy on perinatal outcomes in a Bangladeshi cohort. *Epidemiology*, 27 (2):1, 2015. doi: 10.1097/EDE.0000000000000416.

Koufos, A; Grundy, P; Morgan, K; Aleck, K A; Hadro, T; Lampkin, B C; Kalbakji, A, and Cavenee, W K. Familial Wiedemann-Beckwith syndrome and a second Wilms tumor locus both map to 11p15.5. *American Journal of Human Genetics*, 44(5): 711–9, 1989.

Kuo, C-C; Howard, B V; Umans, J G; Gribble, M O; Best, L G; Francesconi, K A; Goessler, W; Lee, E; Guallar, E, and Navas-Acien, A. Arsenic exposure, arsenic metabolism, and incident diabetes in the Strong Heart Study. *Diabetes Care*, 38(4): 620–7, 2015. doi: 10.2337/dc14-1641.

Laine, J E; Bailey, K A; Rubio-Andrade, M; Olshan, A F; Smeester, L; Drobná, Z; Herring, A H; Stýblo, M; García-Vargas, G G, and Fry, R C. Maternal arsenic

exposure, arsenic methylation efficiency, and birth outcomes in the Biomarkers of Exposure to ARsenic (BEAR) Pregnancy cohort in Mexico. *Environmental Health Perspectives*, 123(2):186–192, 2015. doi: 10.1289/ehp.1307476.

Lim, J C and Donaldson, P J. Focus on molecules: the cystine/glutamate exchanger (System xc). *Experimental Eye Research*, 92(3):162–163, 2011. doi: 10.1016/j.exer.2010.05.007.

Lindberg, A-L; Rahman, M; Persson, L, and Vahter, M. The risk of arsenic induced skin lesions in Bangladeshi men and women is affected by arsenic metabolism and the age at first exposure. *Toxicology and Applied Pharmacology*, 230(1):9–16, 2008. doi: 10.1016/j.taap.2008.02.001.

López-Carrillo, L; Hernández-Ramírez, R U; Gandolfi, A J; Ornelas-Aguirre, J M; Torres-Sánchez, L, and Cebrian, M E. Arsenic methylation capacity is associated with breast cancer in northern Mexico. *Toxicology and Applied Pharmacology*, 280 (1):53–59, 2014. doi: 10.1016/j.taap.2014.07.013.

Melak, D; Ferreccio, C; Kalman, D; Parra, R; Acevedo, J; Pérez, L; Cortés, S; Smith, A H; Yuan, Y; Liaw, J, and Steinmaus, C. Arsenic methylation and lung and bladder cancer in a case-control study in northern Chile. *Toxicology and Applied Pharmacology*, 274(2):225–31, 2014. doi: 10.1016/j.taap.2013.11.014.

Milton, A H; Hussain, S; Akter, S; Rahman, M; Mouly, T A, and Mitchell, K. A review of the effects of chronic arsenic exposure on adverse pregnancy outcomes. *International Journal of Environmental Research and Public Health*, 14(6), 2017. doi: 10.3390/ijerph14060556.

Odewole, O A; Williamson, R S; Zakai, N A; Berry, R J; Judd, S E; Qi, Y P; Adedinsewo, D A, and Oakley, G P. Near-elimination of folate-deficiency anemia by mandatory folic acid fortification in older US adults: reasons for geographic and racial differences in stroke study 2003-2007. *American Journal of Clinical Nutrition*, 98(4): 1042–7, 2013. doi: 10.3945/ajcn.113.059683.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Sanchez, T; van Geen, A; Mey, J L; Siddique, A B; Shahriar, H; Uddin, M N; Islam, T; Balac, O; Ilievski, V; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folic acid and creatine as therapeutic approaches to lower blood arsenic: a randomized controlled trial. *Environmental Health Perspectives*, 123(12):1294–1301, 2015a. doi: 10.1289/ehp.1409396.

Peters, B A; Hall, M N; Liu, X; Parvez, F; Siddique, A B; Shahriar, H; Uddin, M Nasir; Islam, T; Ilievski, V; Graziano, J H, and Gamble, M V. Low-dose creatine supplementation lowers plasma guanidinoacetate, but not plasma homocysteine, in a double-blind, randomized, placebo-controlled trial. *Journal of Nutrition*, 145(10): 2245–52, 2015b. doi: 10.3945/jn.115.216739.

Pilsner, J R; Liu, X; Ahsan, H; Ilievski, V; Slavkovich, V; Levy, D; Factor-Litvak, P; Graziano, J H, and Gamble, M V. Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. *Environmental Health Perspectives*, 117(2):254–60, 2009. doi: 10.1289/ehp.11872.

Pu, Y-S; Yang, S-M; Huang, Y-K; Chung, C-J; Huang, S K; Chiu, A W-H; Yang, M-H; Chen, C-J, and Hsueh, Y-M. Urinary arsenic profile affects the risk of urothelial carcinoma even at low arsenic exposure. *Toxicology and Applied Pharmacology*, 218 (2):99–106, 2007. doi: 10.1016/j.taap.2006.09.021.

Roh, T; Steinmaus, C; Marshall, G; Ferreccio, C; Liaw, J, and Smith, A H. Age at exposure to arsenic in water and mortality 3040 years after exposure cessation. *American Journal of Epidemiology*, 187(11):2297–2305, 2018. doi: 10.1093/aje/kwy159.

Rojas, D; Rager, J E; Smeester, L; Bailey, K A; Drobná, Z; Rubio-Andrade, M; Stýblo, M; García-Vargas, G, and Fry, R C. Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. *Toxicological Sciences*, 143(1):97–106, 2015. doi: 10.1093/toxsci/kfu210.

Schittek, B and Sinnberg, T. Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis. *Molecular Cancer*, 13:231, 2014. doi: 10.1186/1476-4598-13-231.

Smith, A H; Marshall, G; Roh, T; Ferreccio, C; Liaw, J, and Steinmaus, C. Lung, bladder, and kidney cancer mortality 40 years after arsenic exposure reduction. *Journal of the National Cancer Institute*, 110(3):241–249, 2018. doi: 10.1093/jnci/djx201.

Steinmaus, C; Bates, M N.; Yuan, Y; Kalman, D; Atallah, R; Rey, O A; Biggs, M L; Hopenhayn, C; Moore, L E; Hoang, B K, and Smith, A H. Arsenic methylation and bladder cancer risk in case-control studies in Argentina and the United States. *Journal of Occupational and Environmental Medicine*, 48(5):478–488, 2006. doi: 10.1097/01.jom.0000200982.28276.70.

Steinmaus, C; Yuan, Y; Kalman, D; Rey, O A; Skibola, C F; Dauphine, D; Basu, A; Porter, K E; Hubbard, A; Bates, M N; Smith, M T, and Smith, A H. Individual differences in arsenic metabolism and lung cancer in a case-control study in Cordoba, Argentina. *Toxicology and Applied Pharmacology*, 247(2):138–45, 2010. doi: 10.1016/j.taap.2010.06.006.

Steinmaus, C; Ferreccio, C; Acevedo, J; Balmes, J R; Liaw, J; Troncoso, P; Dauphiné, C C; Nardone, A, and Smith, A H. High risks of lung disease associated with early-life and moderate lifetime arsenic exposure in northern Chile. *Toxicology and Applied Pharmacology*, 313:10–15, 2016. doi: 10.1016/j.taap.2016.10.006.

Tseng, C-H; Huang, Y-K Y-L; Chung, C-J; Yang, M-H; Chen, C-J, and Hsueh, Y-M. Arsenic exposure, urinary arsenic speciation, and peripheral vascular disease in blackfoot disease-hyperendemic villages in Taiwan. *Toxicology and Applied Pharmacology*, 206(3):299–308, 2005. doi: 10.1016/j.taap.2004.11.022.

Wang, L-X; Li, Y, and Chen, G-Z. Network-based co-expression analysis for exploring the potential diagnostic biomarkers of metastatic melanoma. *PloS one*, 13(1): e0190447, 2018. doi: 10.1371/journal.pone.0190447.

Wardlaw, T; Blanc, A; Zupan, J; Ahman, E, and Åhman, E. Low birthweight: country, regional, and global estimates. Technical report, New York, NY, 2004.

Wu, M-M; Chiou, H-Y; Hsueh, Y-M; Hong, C-T; Su, C-L; Chang, S-F; Huang, W-L; Wang, H-T; Wang, Y-H; Hsieh, Y-C, and Chen, C-J. Effect of plasma homocysteine level and urinary monomethylarsonic acid on the risk of arsenic-associated carotid atherosclerosis. *Toxicology and Applied Pharmacology*, 216:168–175, 2006. doi: 10.1016/j.taap.2006.05.005.

Yu, R C; Hsu, K-H; Chen, C-J, and Froines, J R. Arsenic methylation capacity and skin cancer. *Cancer Epidemiology, Biomarkers and Prevention*, 9:1259–1262, 2000.

Zeisel, S H. Importance of methyl donors during reproduction. *American Journal of Clinical Nutrition*, 89(2), 2009. doi: 10.3945/ajcn.2008.26811D.

Zhong, Q; Cui, Y; Wu, H; Niu, Q; Lu, X; Wang, L, and Huang, F. Association of maternal arsenic exposure with birth size: a systematic review and meta-analysis. *Environmental Toxicology and Pharmacology*, 69:129–136, 2019. doi: 10.1016/j.etap.2019.04.007.

Zimmerman, S and Lu, C. Defeating Anemia: 2015 Year in Review. Technical report, Food Fortification Initiative, Atlanta, GA, 2015.