Prenatal Arsenic Exposure and Birth Outcomes in a Longitudinal Birth Cohort in Bangladesh

PRENATAL ARSENIC EXPOSURE AND BIRTH OUTCMES IN A LONGITUDINAL BIRTH COHORT IN BANGLADESH

MOHAMMAD RAHMAN

A Dissertation Submitted to the Faculty of

The Harvard T.H. Chan School of Public Health

in Partial Fulfillment of the Requirements

for the Degree of *Doctor of Science*

in the *Department of Environmental Health*

Harvard University

Boston, Massachusetts.

May 2017

Dissertation Advisor: Dr. David C. Christiani Mohammad Rahman

PRENATAL ARSENIC EXPOSURE AND BIRTH OUTCMES IN A LONGITUDINAL BIRTH COHORT IN BANGLADESH

Abstract

This research provides an in-depth investigation into the associations between prenatal arsenic exposure and adverse birth outcomes with a view to mechanistically explain these associations by exploring miRNA expression profiles in the placenta. The first part of the dissertation investigates the effects of prenatal arsenic exposure, adolescent marriage, and pregnancy weight gain on preterm birth in Bangladesh. Because these factors of environmental, societal and maternal health origin often coexist and are highly prevalent in the South and South-East Asia, including Bangladesh, findings from this research may provide insights into effective intervention strategies for reducing preterm birth burden in this region.

The second part of the dissertation critically evaluates the causal association between prenatal arsenic exposure and birthweight in relation to shortening of gestation and intrauterine growth restriction- the two main causal processes of low birthweight. Using quantile causal mediation analysis approach, we estimated pathway specific effects of arsenic exposure on birthweight and evaluated whether the susceptibility of arsenic exposure varies by infant birth sizes.

In the third chapter we delved into identifying placental microRNA markers that regulate birthweight in presence to environmental arsenic exposure. Using epigenome-wide approach, we identified that placenta-derived miRNAs not only regulate birthweight by determining the length of gestation but also modulating the susceptibility of environmental toxicants on fetal growth.

This dissertation presents a more complete understanding of the associations between prenatal arsenic exposure and birthweight by contributing new knowledge to explain underlying mechanisms of this complex exposure-outcome relation.

TABLE OF CONTENTS

List of figures with captions

List of tables with captions

Dedicated to my parents,
Abdur Rahman and Ferdousi Rahman,
and to my wife,
Saima Islam,
for your unconditional love, support, and sacrifice.

Acknowledgements

I would like to express my deep appreciation and gratitude to my advisor, Dr. David C. Christiani, for the patient guidance and mentorship he provided to me. Without his unwavering support, this work would not have been possible. I am forever grateful for his unparalleled intellect, dedication, and motivation to my professional and academic growth. I am truly fortunate to have had the opportunity to work with him.

I would also like to thank my dissertation committee members, Dr. Russ Hauser, Dr. Andrea Baccarelli, and Dr. Liming Liang, for their guidance, through-provoking suggestions, and the dedicated mentorship that each of them offered to me over the years. In a similar vein, I would like to recognize Dr. Molly L. Kile, Dr. Maitreyi Mazumdar, Dr. Linda Valeri, Dr. Ema Rodrigues, and Li Su for the contributions that each of them made to my intellectual growth during my years of study at the Harvard TH Chan School of Public Health. A special thanks to my colleagues at the Dhaka Community Hospital Trust, in Dhaka, Bangladesh for their hard work and commitment to this research. In addition, this work was generously supported by the National Institute of Environmental Health Sciences (NIEHS) of the National Institutes of Health (NIH) grant # R01ES015533.

Finally, I would be remiss if I did not acknowledge the innumerable sacrifices made by my parents, Abdur Rahman and Ferdousi Rahman, and my wife Saima Islam in encouraging and supporting me while I pursued this final degree. Lastly, I am thankful to my son, Ibrahim, for the enormous joy and laughter that he brought to my life during the final year of this work.

Introduction

Arsenic is a ubiquitous metalloid found in various chemical forms in soil, ground water and foods. The environmental exposure to inorganic arsenic occurs largely through ingestion of arsenic-contaminated drinking water (Vahter 2008). Millions of people world-wide are being exposed to arsenic through contaminated drinking water at levels higher than the World Health Organization (WHO) standard of 10 μg/L (Mandal and Suzuki 2002), including approximately 56 million people in Bangladesh (Smith, Lingas, and Rahman 2000) and 13 million in the United States (USEPA 2009).

Inorganic arsenic is a class 1 human carcinogen, and has been linked with carcinoma of skin, lungs, urinary bladder, and kidney (IARC 2004, Celik et al. 2008, Chen et al. 1992, Mink et al. 2008, Vahidnia, van der Voet, and de Wolff 2007, Yu, Liao, and Chai 2006). Among pregnant mothers, arsenic can cross the placental barrier and accumulate in fetal tissues.(Concha, Nermell, and Vahter 1998) Several experimental studies have documented the reproductive and developmental toxicities of arsenic. Evidence from observational studies also linked prenatal arsenic exposure with spontaneous abortion,(Rahman et al. 2010, Milton et al. 2005, He et al. 2007), intrauterine growth restriction(Thomas et al. 2015, Kippler et al. 2012, Llanos and Ronco 2009, Claus Henn et al. 2016), and fetal,(Milton et al. 2005, von Ehrenstein et al. 2006) and neonatal death.(Rahman et al. 2007, Hopenhayn-Rich et al. 2000) However, the epidemiological evidence of the associations between prenatal arsenic exposure and preterm birth and birthweight is less consistent. There also exist significant research gaps in addressing the mechanisms how prenatal arsenic exposure might relate to these adverse birth outcomes.

Arsenic Toxicokinetics

Inorganic arsenic in the form of arsenate (As^{V}) and arsenite (As^{III}) are well absorbed by both the oral and inhalation routes (ATSDR 2007). Absorption by the dermal route has not been well characterized, but is low compared to other routes. Once absorbed, the biotransformation process of inorganic arsenic occurs through a series of cycles of reduction, oxidation, and methylation reactions. The pentavalent arsenate (As^{V}) is reduced to trivalent arsenite (As^{III}), using glutathione and thioredoxin as electron donors, followed by methylation to monomethylarsonic acid (MMA) and then dimethylarsinic acid (DMA) by a series of reactions via one-carbon metabolism using S-Adenosyl methionine (SAM), the universal methyl donor in the body (ATSDR 2007, Vahter 2008).

Approximately 75% of the ingested inorganic arsenic is eliminated shortly in urine as a mixture of 10–30% inorganic arsenic, 10–20% MMA^{V}, and 60–70% DMA^{V} (Vahter 2000). However, the trivalent intermediate metabolites, mainly inorganic As^{III} and MMA^{III}, are much more reactive and have longer half-life in tissues due to their affinity towards sulfhydryl (SH) group. The methylation of arsenic is induced during pregnancy due to the upregulation of one-carbon metabolism to aid placental and fetal tissue development (Concha, Nermell, and Vahter 1998, Hopenhayn et al. 2003). Both inorganic arsenic and its metabolites, MMA and DMA can cross the placenta (Concha, Nermell, and Vahter 1998), facilitated by phosphate transporters and the placental glucose transporters, GLUT1 and GLUT4 (Rosen and Liu 2009). Arsenic can also accumulate in the placental and fetal tissues and induce toxic effects via generating oxygen-derived free radicals (Ahmed et al. 2011) or impairing nutrient transport to the fetus (Lindgren et al. 1984).

Arsenic exposure and the epigenome

Epigenetics is related to stable and heritable patterns of gene expression and genome function that do not involve changes in DNA sequence (Baccarelli and Bollati 2009). Epigenetic processes, such as DNA methylation, histone modification, and miRNA expression, are major regulators of gene expression, and can change genome function under exposure to environmental toxins (Hubaux et al. 2013).

The arsenic-induced epigenetic alterations are mainly derived from deprivation of the cellular pool of methyl (−CH3) residue secondary to the increased demand for SAM during arsenic biotransformation process. Arsenic exposure also interferes the function of DNA methyltransferase, the enzyme responsible for transferring methyl residue from SAM to CpG dinucleotide. As a result, chronic exposure to inorganic arsenic can lead to global hypomethylation and changes in chromatin structure (Intarasunanont et al. 2012, Hubaux et al. 2013). These changes in epigenetic signatures have the potential to greatly impact cellular homeostasis (Haluskova 2010).

The changes in DNA methylation patterns induced by arsenic may lead to the alteration of microRNA (miRNA) expression profiles in a tissue-specific manner. Indeed, there is accumulating evidence from experimental and observational studies that inorganic arsenic can alter gene expression by changing miRNA expression profiles (Baccarelli and Bollati 2009, Ren et al. 2011, Marsit, Eddy, and Kelsey 2006, Rager et al. 2013).

miRNAs are small (18-22 nucleotide in length) non-coding RNAs that regulate gene expression at the post-transcriptional level. They originate from genome-encoded primary transcripts (pri-miRNAs) forming distinctive hairpin structures that are cleaved subsequently by the ribonuclease Drosha (Denli et al. 2004) and Dicer (Bernstein et al. 2001) to imperfect ~22-nt duplexes. One strand of the duplex is incorporated into RNA-induced silencing complex (RISC)

and becomes a functional miRNA (Hutvagner and Zamore 2002). Binding with 3' untranslated region (3'-UTR) of messenger RNA (mRNA) transcript, miRNA essentially regulates 30-60% of all protein-coding genes in human body by promoting mRNA degradation or translational repression (Friedman et al. 2009, Baek et al. 2008). Although the biological significance of many of these miRNAs remains to be elucidated, functional studies have implicated their regulatory actions on cell-cycle progression, apoptosis, inflammation, immune regulation, cell proliferation, and cellular differentiation (Rager et al. 2013, Hagen and Lai 2008).

The placental epigenome and health outcomes

The placenta plays a critical role in maintaining the intrauterine environment in response to extrauterine factors, both maternal and environmental origins, by facilitating fetal supply of nutrients and oxygen, secreting hormones, and, detoxifying potentially harmful maternal factors and environmental toxins.(Gude et al. 2004) In other words, to a large extend the placenta determines the environment to which the growing fetus is exposed. Thus, the placenta is likely to show the strongest evidence of any environmental insult during pregnancy, and that mapping the placental epigenetic profiles may provides a unique opportunity to understand placental responsiveness to the environment.

Extensive epigenetic programming occurs between fertilization and implantation, and many tissue-specific marks are set during the early stages of differentiation. However, epigenetic reprogramming continues to occur throughout prenatal development as well as postnatal and early adulthood period.(Yuen et al. 2011). Thus, epigenetic variation may represent the strongest candidate mechanism linking prenatal exposure with perinatal and adult-life health outcomes (Sundrani et al. 2013).

Overview of Thesis

The objective of this dissertation is to contribute to the ongoing research on prenatal arsenic exposure and birth outcomes with a view to mechanistically explain these associations in relation to miRNA expression profiles in the placenta. The first part of the dissertation investigated the effect of prenatal arsenic exposure on preterm birth, while concurrently assessing the influence of coexisting social (e.g. adolescent marriage) and maternal health factors (e.g. pregnancy weight gain) to this risk. Because these factors often coexist and are highly prevalent in the South and South-East Asia, including Bangladesh, findings from this research may provide insights into effective intervention strategies for reducing preterm birth burden in this region.

In Chapter two, we critically evaluated the causal association between prenatal arsenic exposure and birthweight in relation to shortening of gestation and intrauterine growth restriction- the two main causal processes of low birthweight. Using quantile causal mediation analysis approach, we estimated the pathway specific effects of arsenic on birthweight across birthweight quantiles. Our results showed that the susceptibility of arsenic exposure varies by infant birth sizes, putting smaller infants at a higher risk.

In Chapter three, we delved into identifying placental epigenetic markers that are associated with birthweight with a view to understand the epigenetic regulation of birthweight in relation to environmental arsenic exposure. In a two-stage study, we screened 754 human miRNAs in a discovery cohort of 154 participants sampled from both tails of the birthweight distribution using TaqMan qRT-PCR miRNA Array and then conformed expression of 49 top-hit miRNAs in a replication cohort of 364 participants randomly selected from the full study cohort using probe-based qRT-PCR assay. Using causal mediation analysis approach, we identified

placenta-derived miRNAs that regulate birthweight not only by determining the length of gestation but also modulating the susceptibility of environmental toxicants on fetal growth.

This dissertation aims ultimately to supplement our current knowledge of the associations between prenatal arsenic exposure and birth outcomes by contributing new knowledge to explain the underlying mechanisms of these complex exposure-outcome relations.

Bibliography

Ahmed, S., S. Mahabbat-e Khoda, R. S. Rekha, R. M. Gardner, S. S. Ameer, S. Moore, E. C. Ekstrom, M. Vahter, and R. Raqib. 2011. "Arsenic-associated oxidative stress, inflammation, and immune disruption in human placenta and cord blood." *Environ Health Perspect* 119 (2):258-64. doi: 10.1289/ehp.1002086.

ATSDR. 2007. Agency for Toxic Substance and Disease Registry: Tox guide for Arsenic.

Baccarelli, A., and V. Bollati. 2009. "Epigenetics and environmental chemicals." *Curr Opin Pediatr* 21 (2):243-51.

Baek, D., J. Villen, C. Shin, F. D. Camargo, S. P. Gygi, and D. P. Bartel. 2008. "The impact of microRNAs on protein output." *Nature* 455 (7209):64-71. doi: 10.1038/nature07242.

Bernstein, E., A. A. Caudy, S. M. Hammond, and G. J. Hannon. 2001. "Role for a bidentate ribonuclease in the initiation step of RNA interference." *Nature* 409 (6818):363-6. doi: 10.1038/35053110.

Celik, I., L. Gallicchio, K. Boyd, T. K. Lam, G. Matanoski, X. Tao, M. Shiels, E. Hammond, L. Chen, K. A. Robinson, L. E. Caulfield, J. G. Herman, E. Guallar, and A. J. Alberg. 2008. "Arsenic in drinking water and lung cancer: a systematic review." *Environ Res* 108 (1):48-55. doi: 10.1016/j.envres.2008.04.001.

Chen, C. J., C. W. Chen, M. M. Wu, and T. L. Kuo. 1992. "Cancer potential in liver, lung, bladder and kidney due to ingested inorganic arsenic in drinking water." *Br J Cancer* 66 (5):888-92.

Claus Henn, B., A. S. Ettinger, M. R. Hopkins, R. Jim, C. Amarasiriwardena, D. C. Christiani, B. A. Coull, D. C. Bellinger, and R. O. Wright. 2016. "Prenatal Arsenic Exposure and Birth Outcomes among a Population Residing near a Mining-Related Superfund Site." *Environ Health Perspect* 124 (8):1308-15. doi: 10.1289/ehp.1510070.

Concha, G., B. Nermell, and M. V. Vahter. 1998. "Metabolism of inorganic arsenic in children with chronic high arsenic exposure in northern Argentina." *Environ Health Perspect* 106 (6):355-9.

Denli, A. M., B. B. Tops, R. H. Plasterk, R. F. Ketting, and G. J. Hannon. 2004. "Processing of primary microRNAs by the Microprocessor complex." *Nature* 432 (7014):231-5. doi: 10.1038/nature03049.

Friedman, R. C., K. K. Farh, C. B. Burge, and D. P. Bartel. 2009. "Most mammalian mRNAs are conserved targets of microRNAs." *Genome Res* 19 (1):92-105. doi: 10.1101/gr.082701.108.

Gude, N. M., C. T. Roberts, B. Kalionis, and R. G. King. 2004. "Growth and function of the normal human placenta." *Thromb Res* 114 (5-6):397-407. doi: 10.1016/j.thromres.2004.06.038.

Hagen, J. W., and E. C. Lai. 2008. "microRNA control of cell-cell signaling during development and disease." *Cell Cycle* 7 (15):2327-32.

Haluskova, J. 2010. "Epigenetic studies in human diseases." *Folia Biol (Praha)* 56 (3):83-96.

He, W., R. J. Greenwell, D. M. Brooks, L. Calderon-Garciduenas, H. D. Beall, and J. D. Coffin. 2007. "Arsenic exposure in pregnant mice disrupts placental vasculogenesis and causes spontaneous abortion." *Toxicol Sci* 99 (1):244-53. doi: 10.1093/toxsci/kfm162.

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

Hopenhayn-Rich, C., S. R. Browning, I. Hertz-Picciotto, C. Ferreccio, C. Peralta, and H. Gibb. 2000. "Chronic arsenic exposure and risk of infant mortality in two areas of Chile." *Environ Health Perspect* 108 (7):667-73.

Hubaux, R., D. D. Becker-Santos, K. S. Enfield, D. Rowbotham, S. Lam, W. L. Lam, and V. D. Martinez. 2013. "Molecular features in arsenic-induced lung tumors." *Mol Cancer* 12:20. doi: 10.1186/1476-4598-12-20.

Hutvagner, G., and P. D. Zamore. 2002. "A microRNA in a multiple-turnover RNAi enzyme complex." *Science* 297 (5589):2056-60. doi: 10.1126/science.1073827.

IARC, International Agency for Research on Cancer, Lyon, France. 2004. "Some Drinking Water Disinfectants and Contaminants, Including Arsenic."

Intarasunanont, P., P. Navasumrit, S. Waraprasit, K. Chaisatra, W. A. Suk, C. Mahidol, and M. Ruchirawat. 2012. "Effects of arsenic exposure on DNA methylation in cord blood samples from newborn babies and in a human lymphoblast cell line." *Environ Health* 11:31. doi: 10.1186/1476-069X-11-31.

Kippler, M., Y. Wagatsuma, A. Rahman, B. Nermell, L. A. Persson, R. Raqib, and M. Vahter. 2012. "Environmental exposure to arsenic and cadmium during pregnancy and fetal size: a longitudinal study in rural Bangladesh." *Reprod Toxicol* 34 (4):504-11. doi: 10.1016/j.reprotox.2012.08.002.

Lindgren, A., B. R. Danielsson, L. Dencker, and M. Vahter. 1984. "Embryotoxicity of arsenite and arsenate: distribution in pregnant mice and monkeys and effects on embryonic cells in vitro." *Acta Pharmacol Toxicol (Copenh)* 54 (4):311-20.

Llanos, M. N., and A. M. Ronco. 2009. "Fetal growth restriction is related to placental levels of cadmium, lead and arsenic but not with antioxidant activities." *Reprod Toxicol* 27 (1):88-92. doi: 10.1016/j.reprotox.2008.11.057.

Mandal, B. K., and K. T. Suzuki. 2002. "Arsenic round the world: a review." *Talanta* 58 (1):201-35.

Marsit, C. J., K. Eddy, and K. T. Kelsey. 2006. "MicroRNA responses to cellular stress." *Cancer Res* 66 (22):10843-8. doi: 10.1158/0008-5472.can-06-1894.

Milton, A. H., W. Smith, B. Rahman, Z. Hasan, U. Kulsum, K. Dear, M. Rakibuddin, and A. Ali. 2005. "Chronic arsenic exposure and adverse pregnancy outcomes in bangladesh." *Epidemiology* 16 (1):82-6.

Mink, P. J., D. D. Alexander, L. M. Barraj, M. A. Kelsh, and J. S. Tsuji. 2008. "Low-level arsenic exposure in drinking water and bladder cancer: a review and meta-analysis." *Regul Toxicol Pharmacol* 52 (3):299-310. doi: 10.1016/j.yrtph.2008.08.010.

Rager, J. E., K. A. Bailey, L. Smeester, S. K. Miller, J. S. Parker, J. E. Laine, Z. Drobna, J. Currier, C. Douillet, A. F. Olshan, M. Rubio-Andrade, M. Styblo, G. Garcia-Vargas, and R. C. Fry. 2013. "Prenatal arsenic exposure and the epigenome: Altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood." *Environ Mol Mutagen*. doi: 10.1002/em.21842.

Rahman, A., L. A. Persson, B. Nermell, S. El Arifeen, E. C. Ekstrom, A. H. Smith, and M. Vahter. 2010. "Arsenic exposure and risk of spontaneous abortion, stillbirth, and infant mortality." *Epidemiology* 21 (6):797-804. doi: 10.1097/EDE.0b013e3181f56a0d.

Rahman, A., M. Vahter, E. C. Ekstrom, M. Rahman, A. H. Golam Mustafa, M. A. Wahed, M. Yunus, and L. A. Persson. 2007. "Association of arsenic exposure during pregnancy with fetal loss and infant death: a cohort study in Bangladesh." *Am J Epidemiol* 165 (12):1389-96. doi: 10.1093/aje/kwm025.

Ren, X., C. M. McHale, C. F. Skibola, A. H. Smith, M. T. Smith, and L. Zhang. 2011. "An emerging role for epigenetic dysregulation in arsenic toxicity and carcinogenesis." *Environ Health Perspect* 119 (1):11-9. doi: 10.1289/ehp.1002114.

Rosen, B. P., and Z. Liu. 2009. "Transport pathways for arsenic and selenium: a minireview." *Environ Int* 35 (3):512-5. doi: 10.1016/j.envint.2008.07.023.

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

Sundrani, D. P., U. S. Reddy, A. A. Joshi, S. S. Mehendale, P. M. Chavan-Gautam, A. A. Hardikar, G. R. Chandak, and S. R. Joshi. 2013. "Differential placental methylation and expression of VEGF, FLT-1 and KDR genes in human term and preterm preeclampsia." *Clin Epigenetics* 5 (1):6. doi: 10.1186/1868-7083-5-6.

Thomas, S., T. E. Arbuckle, M. Fisher, W. D. Fraser, A. Ettinger, and W. King. 2015. "Metals exposure and risk of small-for-gestational age birth in a Canadian birth cohort: The MIREC study." *Environ Res* 140:430-9. doi: 10.1016/j.envres.2015.04.018.

USEPA. 2009. "United States Environmental Protection Agency: Fact Sheet: drinking water standard for arsenic." (EPI815-R-00-015).

Vahidnia, A., G. B. van der Voet, and F. A. de Wolff. 2007. "Arsenic neurotoxicity--a review." *Hum Exp Toxicol* 26 (10):823-32. doi: 10.1177/0960327107084539.

Vahter, M. 2000. "Genetic polymorphism in the biotransformation of inorganic arsenic and its role in toxicity." *Toxicol Lett* 112-113:209-17.

Vahter, M. 2008. "Health effects of early life exposure to arsenic." *Basic Clin Pharmacol Toxicol* 102 (2):204-11. doi: 10.1111/j.1742-7843.2007.00168.x.

von Ehrenstein, O. S., D. N. Guha Mazumder, M. Hira-Smith, N. Ghosh, Y. Yuan, G. Windham, A. Ghosh, R. Haque, S. Lahiri, D. Kalman, S. Das, and A. H. Smith. 2006. "Pregnancy outcomes, infant mortality, and arsenic in drinking water in West Bengal, India." *Am J Epidemiol* 163 (7):662-9. doi: 10.1093/aje/kwj089.

Yu, H. S., W. T. Liao, and C. Y. Chai. 2006. "Arsenic carcinogenesis in the skin." *J Biomed Sci* 13 (5):657-66. doi: 10.1007/s11373-006-9092-8.

Yuen, R. K., S. M. Neumann, A. K. Fok, M. S. Penaherrera, D. E. McFadden, W. P. Robinson, and M. S. Kobor. 2011. "Extensive epigenetic reprogramming in human somatic tissues between fetus and adult." *Epigenetics Chromatin* 4:7. doi: 10.1186/1756-8935-4-7.

1. Chapter 1 - Prenatal arsenic exposure, adolescent marriage, and pregnancy weight gain: Associations with preterm birth in Bangladesh

Mohammad Rahman[1], Molly L. Kile[2], Ema G Rodrigues[1], Linda Valeri[3], Maitreyi Mazumdar[1,4], Russ Hauser[1], Andrea Baccarelli[5], Liming Liang[1], Golam Mostofa[6], Quazi Quamruzzaman[6], Mahmudur Rahman[6], David C. Christiani[1, 3]

[1]Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

[2] Oregon State University, College of Public Health and Human Sciences, Corvallis, Oregon, USA

[3]Department of Psychiatry (Biostatistics), McLean Hospital, Belmont, MA, USA

[4]Department of Neurology, Boston Children's Hospital, Boston, MA, USA

[5]Columbia University, Mailman School of Public Health, New York, USA

[6]Dhaka Community Hospital Trust, Dhaka, Bangladesh

Abstract

Background: Preterm birth is a disease of multifactorial etiologies that has environmental, social, and maternal health components. Often, these factors coexist in a given population, and that they may have interactive effects on preterm birth.

Objective: To evaluated the associations between prenatal arsenic exposure, adolescent marriage (marriage <18 years), pregnancy weight gain, and preterm birth in a longitudinal birth cohort in Bangladesh.

Methods: During 2008-2011, 1,613 pregnant women aged ≥16 years were recruited at ≤16 weeks of gestation and followed throughout pregnancy. At the end of follow-up, 1,184 singleton livebirths were recorded. We measured arsenic in drinking water (n=1,182) collected at enrollment and maternal toenails (n=1,104) collected within one-month postpartum using inductively coupled plasma mass spectrometry, determined gestational age using ultrasound, collected demographic information, including self-reported age of marriage using structured questionnaire, and calculated 2nd and 3rd trimester pregnancy weight gain using monthly records.

Results: In adjusted multivariable modified poisson regression models, the risk ratios for preterm birth were 1.12 (95% CI: 1.07−1.18) for a unit change in natural log water arsenic exposure, 2.28 (95%CI: 1.76−2.95) for adolescent marriage, and 0.64 (95% CI: 0.42−0.97) for a unit change in pregnancy weight gain (pound/week). In stratified analysis by adolescent marriage, pregnancy weight gain reduced the risk of preterm birth only among mothers married as adolescents (RR=0.58; 95% CI: 0.37−0.92). Similar associations were observed for maternal toenail arsenic.

Conclusions: Minimizing arsenic exposure during pregnancy and avoiding adolescent marriage can reduce the risk of preterm birth in Bangladesh. Additionally, nutritional support to ensure

adequate weight gain during pregnancy among mothers with a history of adolescent marriage may provide additional benefit.

Introduction

Preterm birth (livebirth <37 weeks of gestation) is the leading cause of neonatal mortality and the second leading cause of mortality among children under five years of age; the surviving infants are at lifelong risk of neurological impairments and chronic lung diseases.(Mwaniki et al. 2012) Worldwide, more than 15 million babies are born preterm each year, and more than a third of them are in South Asia.(Blencowe et al. 2012) In Bangladesh, the incidence of preterm birth is estimated to be 14%,(Blencowe et al. 2012) which is one of the highest in the world, while in rural areas, the estimates are as high as 19-22%.(Baqui et al. 2013, Shah et al. 2014) The higher incidence of preterm birth in Bangladesh is thought to be linked with, among other factors, poor maternal health status during pregnancy, higher prevalences of prenatal arsenic exposure, and adolescent marriage (marriage <18 years old).

Bangladesh has one of the world's highest prevalences of adolescent marriage, where 66% of the women are married before age 18 years, and 29% are married before age 15 years.(UNFPA 2012) Adolescent marriage often forces women into adolescent pregnancy, as adolescent brides have less control over their reproductive decision and are less likely to use contraception prior to first childbirth.(Raj et al. 2009) Adolescent marriage thus, may serve as a good proxy for adolescent pregnancy. This is particularly true in countries like Bangladesh, where there is social expectations for women to get pregnant immediately or soon after marriage and that pregnancy outside marriage is very rare.(Sayem and Nury 2011) Women married as adolescents often share common susceptibility characteristics such as low socio-economic background, and are more likely to have multiple and closely-spaced pregnancies and acquire

perinatal infections, while less likely to utilize maternal health services during pregnancy, which are independent risk factors for preterm birth.(Raj et al. 2009, Raj and Boehmer 2013) Therefore, the association between adolescent marriage and preterm birth could partly be due to unadjusted factors that are related to both adolescent marriage and preterm birth. The biological immaturity of adolescent mothers to sustain pregnancy could also contribute to the risk.

Furthermore, an estimated 30-70 million people in Bangladesh are exposed to arsenic through contaminated drinking water at concentrations higher than the World Health Organization (WHO) guideline value of 10µg/L.(Kinniburgh and Smedley 2001) Arsenic readily can cross placenta (Concha, Nermell, and Vahter 1998, Rudge et al. 2009) and several studies have reported that prenatal arsenic exposure increases the risk of spontaneous abortion (Milton et al. 2005), stillbirth (Milton et al. 2005, von Ehrenstein et al. 2006), infant death (Rahman et al. 2007, Hopenhayn-Rich et al. 2000), and low birth weight (Hopenhayn et al. 2003, Huyck et al. 2007). However, the association between arsenic exposure and preterm birth is inconsistent, which could partly be due to ecological exposure assessment strategies adopted by most studies.

Both adolescent marriage(Scholl and Hediger 1993) and prenatal arsenic exposure(Kile et al. 2016, Kile et al. 2014) impact maternal health status during pregnancy, resulting in inadequate gestational weight gain, which is an independent risk factor for preterm birth. For instance, studies in developing and developed countries reported that low pregnancy weight gain, measured in both total pregnancy weight gain or weekly weight gain during the 2nd and 3rd trimesters, is associated with increased risk of preterm birth.(Han et al. 2011) However, despite the intricate relations among these risk factors of social, environmental and maternal health origin, and that they often coexist in many parts of the world, especially in South and South-East

Asia, most studies focused on these risk factors in isolate with regards to preterm birth, where they might have interactive effects on preterm birth.

In our analysis, we estimated individual as well as interactive effects of prenatal arsenic exposure, adolescent marriage, and pregnancy weight gain on preterm birth among adult mothers in a longitudinal birth cohort in Bangladesh.

Methods

Study participants

We used data from an ongoing longitudinal birth cohort established in Bangladesh during 2008-2011. The details of this study, including details of recruitment and enrollment have been previously reported (Kile et al. 2014). Briefly, women were eligible to participate if they were 18 years or older with an ultrasound confirmed singleton pregnancy of ≤16 weeks' gestation, used a tube well as their primary source of drinking water and had been using the same drinking water source for more than six months, intended to live in her current residence throughout pregnancy. Total 1,613 pregnant women were recruited and followed throughout pregnancy. 1,184 singleton livebirths were recorded at the end of follow-up. Total 99 participants lost to contact before

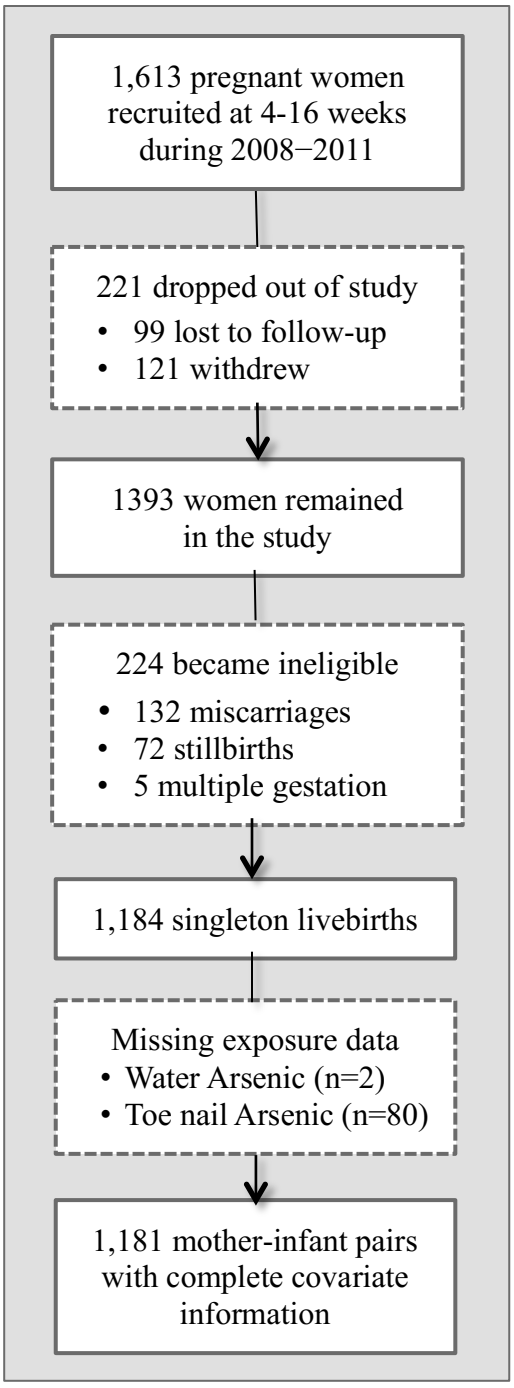

Figure 1.1. Study Participation

delivery, 121 participats withdrew from the study, 132 pregnancies terminated in miscarriage, and 72 fetal deaths and 5 twin births were recorded. Missing data for drinking water and maternal toenail arsenic exposure was recorded for 2 and 80 participants, respectively (Figure-1.1). Women in our sample were similar in respect to chronological age, age at marriage, exposure to secondhand smoke, parity, household income, education, enrollment BMI, and blood hemoglobin level with those lost to follow-up.

All subjects provided written informed consent prior to participation. Participants were informed and counseled on safe drinking water options if their water samples contained arsenic above the Bangladesh standard. Prenatal care and multivitamins were provided to all participants. All protocols were reviewed and approved by the Human Research Committees at Harvard T.H. Chan School of Public Health and Dhaka Community Hospital Trust.

Exposure Assessment:

Adolescent marriage was defined as formal marriage before 18 years of age, which is the legal age of marriage for girls in Bangladesh. Structured questionnaire was used to collect self-reported age at marriage at the time of enrollment. Arsenic was measured in drinking water collected from tubewell that women identified as their primary source of drinking water at the time of enrollemnt. Water samples were collected in 50-ml polypropylene tubes (BD Falcon, BD Bioscience, Bedford, MA) and preserved with Reagent Grade nitric acid (Merck, Germany) to a pH <2. Samples were kept at room temperature prior to analysis by inductively coupled plasma mass spectrometry following US EPA method 200.8 (Environmental Laboratory Services, North Syracuse, New York) with details been previously described (Kile et al. 2014). The average percent recovery of arsenic was 101% (range: 92%-110%). Samples with arsenic concentrations

below the limit of detection (LOD) of 1μg/L were re-assigned a value of half the LOD for statistical analysis.

Arsenic was also measured in maternal toenails collected ≤1-month post-partum. Toenail samples were sonicated in 1% Triton X-100 solution (Sigma-Aldrich, Inc., St. Louis,MO) and rinsed repeatedly with Milli-Q water (Millipore Corporation, Billerica,MA) to remove external contamination before microwave acid digestion using Trace Select Ultra Pure nitric acid (Fischer Scientific, Pittsburgh,PA). Digested samples were diluted with Milli-Q water before analyzing for total arsenic using ICP-MS. The reported arsenic concentrations were blank-corrected and normalized using arsenic concentration of certified human hair reference material (CRM Hair; Shanghi Institute of Nuclear Research, China). Toenail clippings were collected from 1,155 samples, 1,104 of them were used in the analysis after excluding samples with low mass (≤5mg) and/or high relative standard deviation (≥25%). Samples with arsenic concentrations below the LOD ranging from 0.09−0.7ng/L were reassigned half the value of LOD for statistical analysis.

Maternal height and weight was measured at first clinic visit and at monthly house visits following enrollment on adult scale that was calibrated before each measurement and rounded to the nearest 10 oz. We estimated maternal weight gain during the second and third trimesters based on previous studies that reported that preterm birth association is strongest for maternal weight gain during the later trimesters of pregnancy, (Siega-Riz, Adair, and Hobel 1996, Hediger et al. 1989) when weight gain is also relatively higher and linear.(Rasmussen and Yaktine 2009) Basically, we fit a linear regression model with maternal monthly weight measurements as a dependent variable (y) and gestational weeks from 14th week to the last available weight measured prior to delivery as an independent variable (x). The estimated slope of the linear

model reflected pregnancy weight gain during the second and third trimesters as a function of week (pound/week) for each individual.

Outcome and covariates

Preterm birth was defined as livebirth before 37 completed weeks of gestation. Health care workers attended all births to collect detail birth records. Gestational age was determined by ultrasonography by a licensed general practitioner using either (1) gestational sac mean diameter if the pregnancy was between 4-7 weeks or (2) crown-rump length if the pregnancy was between 7-16 weeks.(Hellman et al. 1969, Robinson and Fleming 1975) Other demographic information were collected at the time of enrollment ranging 4-16 weeks of gestation. Maternal hemoglobin concentration was measured in whole blood collected at the time of enrollment. Body mass index (BMI) was categorized based on the World Health Organization (WHO) definitions: underweight, <18·50; normal weight, 18·50–24·99; overweight, ≥25·00.

Statistical Analysis

Drinking water arsenic concentrations were right skewed and subsequently transformed to their natural log (ln). The bivariate and multivariate associations between preterm birth and risk factors were analyzed using Poisson regression with robust variance and risk ratios were estimated.(Zou 2004) Models were adjusted for natural log water arsenic (continuous), adolescent marriage (yes, no), pregnancy weight gain (continuous), maternal age (continuous), education (no formal education, primary, secondary or higher), number of past pregnancy (0, 1, ≥2), BMI categories (<18·50, 18·50–24·99, ≥25·00), and blood hemoglobin level (continuous). We also estimated proportion of preterm birth among women married as adolescents that is attributable to adolescent marriage and factors associated with it (e.g. attributable fraction, AF) using the formula: $\mathrm{AF} = (\mathrm{RR} - 1)/\mathrm{RR}$, where RR is the risk ratio of adolescent marriage.

Standard errors of attributable fraction and population attributable fraction were estimated using delta-method.(Oehlert 1992) Analyses were implemented with STATA (Version 14.0, StataCorp LP, College Station, Texas).

Results

Of the 1,182 singleton livebirths included in our analysis, 260 (22%) were born preterm. Approximately 70% of infants born preterm were born to women who had a history of adolescent marriage. The mean age of marriage was 17.4 years (range:11−29 years); 45% of the women were married before age 18, compared to the national average of 66%.(UNFPA 2012) Mothers married as adolescents were more likely to have low educational attainment, low household income, low BMI, but high chronological age, and multiple past pregnancies. Drinking water arsenic exposure was relatively modest with a median concentration of 2·2µg/L (interquartile range: 0.9µg/L−36µg/L) but spanned a wide range (<LOD– 1400 µg/L). Nearly 22% of the mothers were exposed to drinking water arsenic at concentrations above the national standard of 50 µg/L and contributed to over 35% of all preterm births. Maternal weight gain during the second and third trimesters followed a linear trend (Figure-1.2) with mean 0·80 (interquartile range: 0·55−0·98 pound/week). Other population characteristics and and their association with preterm birth are presented in Table-1.1.

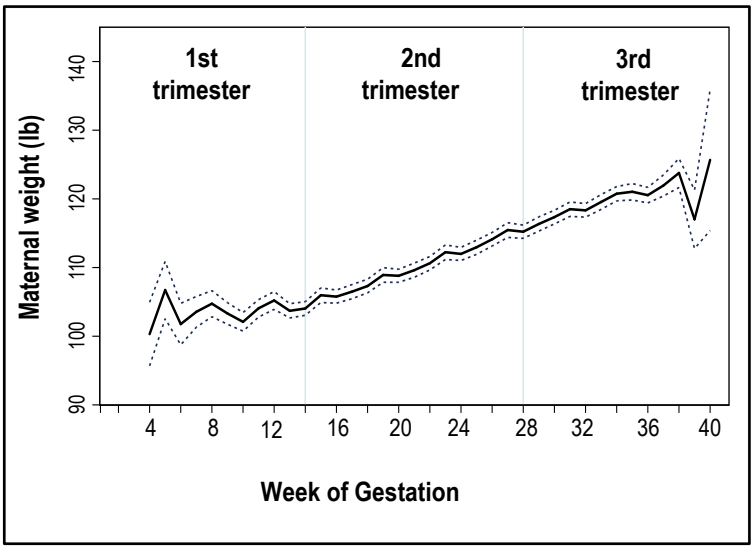

Figure 1.2. Mean body weight by week of gestation. Solid line indicate mean observed body weight and dotted lines indicate one standard deviation above and bellow each mean

Table 1.1- Distribution of Selected Characteristics and their Relations with Preterm Birth

Characteristic	**Preterm Birth** (n = 260)*	**Term Birth** (n = 922)*	**Risk Ratio** (95% CI)	**P-value**
Maternal age (years)	22·97 ± 3·94	22·95 ± 4.29	1·00 (0·98, 1·03)	0·97
Adolescent marriage, n (%)				
Yes (married < 18 year)	181 (70·0)	348 (37·7)	2.86 (2·25, 3.63)	<0·0001
No (married ≥18 year)	79 (30·1)	574 (62·3)	1·00	
2nd & 3rd trimester weight gain (lb/wk)	0·74 ± 0·26	0·81 ± 0·27	0·46 (0·30, 0·70)	<0·0001
Enrollment BMI, n (%)				
≤18·5 kg/m^2	71 (27.0)	265 (28·7)	0·88 (0·69, 1·12)	0·29
>18·5 to ≤25.0 kg/m^2	176 (68.0)	562 (61.0)	1·00	
>25.0 kg/m^2	13 (5·0)	95 (10·3)	0·50 (0·30, 0·85)	0·01
No. of past pregnancies, n (%)				
0	86 (32·8)	390 (42·3)	1·00	
1	82 (31·7)	271 (29·4)	1·30 (0·99, 1·70)	0·06
≥ 2	92 (35·5)	261 (28·3)	1·46 (1·12, 1·89)	0·005
Maternal education, n (%)				
No formal education	49 (18·9)	123 (13·3)	1·00	
Primary education	61 (23·2)	320 (34·7)	0·55 (0·40, 0·77)	<0·0001
Secondary or higher	150 (57·9)	479 (52·0)	0·84 (0·64, 1·10)	0·20
Blood hemoglobin (gm/L)	10·59 ± 1·38	11·21 ± 1·34	0·77 (0·70, 0·83)	<0·0001
Secondhand smoke, n (%)				
Yes	121 (46·9)	374 (40·6)	1·22 (0·99, 1·52)	0·07
No	138 (53·1)	548 (75·4)	1·00	
Natural log water arsenic (μg/L)	2·59 ± 2·23	1·41 ± 2·19	1·19 (1·14, 1·25)	<0·0001

* Values are mean ± Standard Error of the Mean (SEM) except where indicated

Table 1.2 displays the results of multivariate Poisson regression models for the associations between adolescent marriage, prenatal drinking water arsenic exposure, pregnancy weight gain, and preterm birth. Mothers who had a history of adolescent marriage had 128% (risk ratio [RR]=2.28; 95% confidence intervals [CI]: 1.76-2.95) increased risk of preterm birth

compared to mothers with no history of adolescent marriage. For drinking water arsenic, every unit increase in natural log arsenic exposure increased the risk of preterm birth by 12% (RR=1.12; 95% CI: 1.07-1.18) in the whole cohort. In stratified analysis, the risks of preterm birth among mothers with no history of adolescent marriage (RR=1.15; 95% CI: 1.05-1.26) and mothers with a history of adolescent marriage (RR=1.12, 95% CI: 1.05-1.17) remained consistent for drinking water arsenic expsoure. On the contrary, pregnancy weight showed an inverse associations with preterm birth. For instance, every pound increase in maternal weight during the 2nd and 3rd trimesters decreased the risk of preterm birth by 36% (95% CI: 0.42-0.97) in the whole cohort. However, in stratified analysis, the benificial effect of pregnancy weight gain was observed only among mothers who had a a history of adolescent marriage (RR=0.58; 95% CI: 0.37-0.92).

Similarly, when maternal toenail arsenic was used in the modles, the associations between adolescent marriage, prenatal arsenic exposure, pregnancy weight gain, and preterm birth remained consistent, although the effects of toenail arsenic exposure on preterm birth in the statified analysis were not statistically signifiacnt.

Table 1.2- Associations between adolescent marriage, drinking water arsenic exposure, pregnancy weight gain, and preterm birth

Predictors	**Full Cohort** (n=1,182)		**Married as Adults** (n=653)		**Married as Adolescents** (n=529)	
	Risk Ratio (95% CI)	*P*-value	Risk Ratio (95% CI)	*P*-value	Risk Ratio (95% CI)	*P*-value
Adolescent marriage	2.28 (1.76, 2.95)	<0.0001				
Natural log water arsenic exposure (μg/L)	1.12 (1.07,1.18)	<0·0001	1.15 (1.05, 1.26)	0.003	1.12 (1.05, 1.17)	<0.001
Pregnancy weight gain (lb/wk)	0.64 (0.42, 0.97)	0.04	0.86 (0.37, 2.01)	0.72	0.58 (0.37, 0.92)	0.02

Table 1.3- Associations between adolescent marriage, maternal toenail arsenic exposure, pregnancy weight gain, and preterm birth

Predictors	**Full Cohort** (n=1,107)		**Married as Adults** (n=612)		**Married as Adolescents** (n=495)	
	Risk Ratio (95% CI)	*P*-value	Risk Ratio (95% CI)	*P*-value	Risk Ratio (95% CI)	*P*-value
Adolescent marriage	2.44 (1.85, 3.20)	<0.0001				
Natural log toenail arsenic exposure (µg/g)	1.12 (1.02,1.23)	0·02	1.15 (0.96, 1.37)	0.13	1.09 (0.98, 1.22)	0.12
Pregnancy weight gain (lb/wk)	0.58 (0.38, 0.89)	0.01	0.86 (0.36, 2.07)	0.74	0.49 (0.30, 0.80)	0.004

Using methods to simulate the impact of interventions, we estimated the percent of preterm birth burden attributable to adolescent marriage in our cohort. We estimated that 38.0% (95% CI: 26·3-47·9) of preterm birth burden among all mothers and 54.4% (95% CI: 40·9-64·9) of preterm birth burden among mothers with a history of adolescent marriage were attributable to adolescent marriage and factors associated with it. Further, we estimated that by reducing drinking water arsenic exposure from Bangladeshi standard of 50µg/L to the WHO recommended level of 10µg/L, nearly 20.5% (95% CI: 19.3-21.7) of the arsenic-related preterm births could have been avoided.

Discussion

Results from this study demonstrate that adolescent marriage and prenatal arsenic exposure was associated with increased risk of preterm birth. The association between prenatal arsenic exposure and preterm birth showed a dose-response relationship, where every unit increase in natural log arsenic exposure increased the risk of preterm birth by 12%, and that the associations were consistent for arsenic exposure measured in drinking water at the time of

enrollment and in maternal toenails collected within 1-month post-partum, providing evidence for robustness of this finding. On the contrary, maternal 2nd and 3rd trimester weight gain was associated with decreased risk of preterm birth, particularly among mothers with a history of adolescent marriage, suggesting possible scope of intervention among this high risk group by improving perinatal nutrition.

The positive association between adolescent pregnancy and preterm birth has been reported previously.(Chen et al. 2007, Cunnington 2001, Fraser, Brockert, and Ward 1995, Gibbs et al. 2012) These studies examined the effect of young maternal age on adverse birth outcomes in the context of index adolescent pregnancy, perhaps because it has been assumed that the added years between pregnancies reduce obstetric risk. However, some researchers argue that the risk of adverse pregnancy outcomes associated with adolescent pregnancy may continue even in subsequent pregnancies (Jekel et al. 1975) or perhaps, throughout the reproductive life of the mother (Shawky and Milaat 2000). For example, a hospital based case control study in Saudi Arabia reported that women of early teenage marriage (<16 years) experienced higher odds of spontaneous abortion, fetal death, and infant mortality throughout their childbearing period (Shawky and Milaat 2000). This suggests that the health and wellbeing in adolescence can be compromised by adolescent marriage, resulting a long term maternal and child health consequences (Raj et al. 2010).

In addition to adolescent pregnancy, adolescent marriage, can also be considered as a marker for low socioeconomic status, poorer quality and quantity of antenatal care, higher frequency of perinatal infection, physically demanding work, and higher levels of stress, and other adverse psychological factors, which could also independently increase the risk of preterm birth.(Scholl, Hediger, and Belsky 1994, Kramer and Lancaster 2010, UNFPA 2013) Hence, the

observed positive association between adolescent marriage and preterm birth in out cohort could partly be attributed to the list of factors that are related to adolescent marriage. This also explains why we observed higher risk of preterm birth among mothers with a history of adolescent marriage even when they reached their adulthood, as the heightened socioeconomic vulnerability for mothers married as adolescents are likely to persist even during subsequent adulthood pregnancies. In addition, adolescent mothers often suffer essential micronutrient deficiency (Lenders, McElrath, and Scholl 2000) and are also more likely to have repeated and closely spaced pregnancies (Raj et al. 2009), which increases their risk of entering a reproductive cycle with reduced nutritional reserves even when they reach adulthood. As a result, the risk of preterm birth can remains consistently high among this group through out reproductive lives.

The positive associations between prenatal arsenic exposure and preterm birth have been reported previously (Ahmad et al. 2001, Xu et al. 2011, Shi et al. 2015), but few other studies cast doubt upon this relation by reporting null associations (Ahamed et al. 2006, Mukherjee et al. 2005, Myers et al. 2010, Vall et al. 2012). Non-significant positive associations were also reported by two small cross sectional studies in India (Chakraborti et al. 2003, Mukherjee et al. 2005) and a large (n=18,259) prospective study in Taiwan(Yang et al. 2003). It is worth noting that these studies were limited either in sample size or individual-level exposure data, which might explain mixed results among these studies. More recently, Claus Henn et al, reported an inverse association between maternal blood arsenic and gestational age in a rural US cohort (n=622) residing near a mining site after adjusting for metal co-exposure.

The mechanisms leading to arsenic's effect on preterm birth are not well understood, although there is a strong link with infection or inflammation (Goldenberg et al. 2008). Prenatal arsenic exposure has been associated with increased inflammatory markers in newborn cord

blood (Ahmed et al. 2011, Fry et al. 2007). Maternal total urine arsenic was also found associated with altered expression of cord blood microRNAs that are integral to inflammatory and immune pathways (Rager et al. 2013). Abnormal placentation is a known risk factor for preterm birth (Murphy et al. 2006), and inorganic arsenic has been associated with altered angiogenesis leading to dysplastic placental development (He et al. 2007).

The associations between pregnancy weight gain and preterm birth have also been investigated in several studies. While, few studies reported a null association (Marsoosi, Jamal, and Eslamian 2004, Varma 1984), a growing number of studies found an inverse association (Dietz et al. 2006, Han et al. 2011, Schieve et al. 2000, Siega-Riz, Adair, and Hobel 1994, Spinillo et al. 1998). A recent meta-analysis reported that the overall risk of preterm birth was significantly higher among women gaining less total weight gain during pregnancy (overall RR=1·64; 95% CI: 1·62-1·65) or weekly weight (overall RR=1·56; 95% CI: 1·26-1·94) (Han et al. 2011), roughly based on the Institute of Medicine (IOM) guideline (IOM 1990). Other studies reported that later trimester weight gain has stronger inverse association with preterm birth (Siega-Riz, Adair, and Hobel 1994, Scholl and Hediger 1995).

In our analysis, we observed that the inverse association between pregnancy weight gain and preterm birth was modified by mother's prior history of adolescent marriage. As mentioned earlier, the reduced nutritional reserve of mothers with a history of adolescent marriage may exert feto-maternal competitions for nutrients, which has been linked with conditions including smaller placental mass, less placental nutrient transfer, and less uterine and umbilical cord blood transfer predisposing to preterm birth (Wallace et al. 2004). Therefore, ensuring adequate perinatal nutrition among this high-risk group may provide selective advantage in reducing the risk of preterm birth. The current IOM guideline also suggests relatively higher weight gain for

pregnant adolescents to minimize the risk of adverse perinatal outcomes (King 2003, Rasmussen and Yaktine 2009).

As in any observational studies, our study may be limited by unmeasured confounding. While we considered both confounders and predictors of preterm birth, unmeasured confounding by factors including maternal physical activity level and psychological stress during pregnancy, prior history of adverse birth outcomes, and pregnancy complications is possible. Recall bias and potential inaccuracies in reporting maternal age and age at marriage were a possibility, as no reliable birth or marriage registry existed to validate such information.

Our study builds upon existing literature on adolescent marriage, prenatal arsenic exposure, and pregnancy weight gain by bridging the gap among social, environmental, and nutritional determinants of preterm birth. Our use of drinking water collected at enrollment and maternal toenail collected 1-month post partum to measure prenatal arsenic exposure allowed us to reduce the risk of exposure misclassification. Maternal toenail arsenic represents individual's cumulative exposure over the past 9−12 months (Chen, Amarasiriwardena, and Christiani 1999), corresponding the entire pregnancy. Drinking water arsenic, which shows little temporal variability (Slotnick, Meliker, and Nriagu 2006), may serve as an adequate marker for individual's long-term exposure, particularly when measured in main water sources (Sohel et al. 2010). Therefore, our use of both drinking water and maternal toenail arsenic will provide the most representative measure of *in-utero* arsenic exposure. The replicative results between drinking water and maternal toenail arsenic also strengthen our study finding.

Our use of 2nd and 3rd trimester weekly weight gain will also minimize the risk of differential exposure misclassification that often arise with total pregnancy weight gain in relation to preterm birth. Maternal weight gain is not uniform throughout pregnancy; it is

relatively low during the 1st trimester and higher and linear during the 2nd and 3rd trimesters (Rasmussen and Yaktine 2009). Therefore, the total weight gain among women with preterm delivery will be systematically lower compared to women with term delivery because of shorter gestation and disproportionate contribution of lower 1st trimester weight gain (Schieve et al. 2000). Additionally, the prospective nature of this study and a relatively large sample size is another strength. All study subjects received the same level of prenatal care and vitamin supplements with reportedly 99% compliance through our community outreach clinics, which are among the few health service providing centers in that catchment area. The study subjects were ethnically similar, and all were married and non-smokers. Finally, findings of our research provide insight into designing effective intervention strategies for reducing preterm birth burden in Bangladesh and other developing countries with higher prevalences of adolescent marriage and environmental arsenic exposure.

Conclusions

Our results indicate that adolescent marriage, prenatal arsenic exposure, and poor pregnancy weight gain are linked with higher prevalence of preterm birth in Bangladesh. Public health interventions to address societal, economic, and health care issues that are related to adolescent marriage as well as minimizing arsenic exposure during pregnancy may improve perinatal outcomes in Bangladesh. Improving maternal nutritional status during pregnancy can provide additional benefit to the high-risk mothers with a history of adolescent marriage, especially when abolishing adolescent marriage is not feasible in a short time scale.

Bibliography

Ahamed, S., M. K. Sengupta, S. C. Mukherjee, S. Pati, A. Mukherjeel, M. M. Rahman, M. A. Hossain, B. Das, B. Nayakl, A. Pal, A. Zafar, S. Kabir, S. A. Banu, S. Morshed, T. Islam, M. M. Rahman, Q. Quamruzzaman, and D. Chakraborti. 2006. "An eight-year study report on arsenic contamination in groundwater and health effects in Eruani village, Bangladesh and an approach for its mitigation." *J Health Popul Nutr* 24 (2):129-41.

Ahmad, S. A., M. H. Sayed, S. Barua, M. H. Khan, M. H. Faruquee, A. Jalil, S. A. Hadi, and H. K. Talukder. 2001. "Arsenic in drinking water and pregnancy outcomes." *Environ Health Perspect* 109 (6):629-31.

Ahmed, S., S. Mahabbat-e Khoda, R. S. Rekha, R. M. Gardner, S. S. Ameer, S. Moore, E. C. Ekstrom, M. Vahter, and R. Raqib. 2011. "Arsenic-associated oxidative stress, inflammation, and immune disruption in human placenta and cord blood." *Environ Health Perspect* 119 (2):258-64. doi: 10.1289/ehp.1002086.

Baqui, A. H., H. E. Rosen, A. C. Lee, J. A. Applegate, S. El Arifeen, S. M. Rahman, N. Begum, R. Shah, G. L. Darmstadt, and R. E. Black. 2013. "Preterm birth and neonatal mortality in a rural Bangladeshi cohort: implications for health programs." *J Perinatol* 33 (12):977-81. doi: 10.1038/jp.2013.91.

Blencowe, H., S. Cousens, M. Z. Oestergaard, D. Chou, A. B. Moller, R. Narwal, A. Adler, C. Vera Garcia, S. Rohde, L. Say, and J. E. Lawn. 2012. "National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications." *Lancet* 379 (9832):2162-72. doi: 10.1016/S0140-6736(12)60820-4.

Chakraborti, D., S. C. Mukherjee, S. Pati, M. K. Sengupta, M. M. Rahman, U. K. Chowdhury, D. Lodh, C. R. Chanda, A. K. Chakraborti, and G. K. Basu. 2003. "Arsenic groundwater contamination in Middle Ganga Plain, Bihar, India: a future danger?" *Environ Health Perspect* 111 (9):1194-201.

Chen, K. L., C. J. Amarasiriwardena, and D. C. Christiani. 1999. "Determination of total arsenic concentrations in nails by inductively coupled plasma mass spectrometry." *Biol Trace Elem Res* 67 (2):109-25. doi: 10.1007/BF02784067.

Chen, X. K., S. W. Wen, N. Fleming, K. Demissie, G. G. Rhoads, and M. Walker. 2007. "Teenage pregnancy and adverse birth outcomes: a large population based retrospective cohort study." *Int J Epidemiol* 36 (2):368-73. doi: 10.1093/ije/dyl284.

Concha, G., B. Nermell, and M. V. Vahter. 1998. "Metabolism of inorganic arsenic in children with chronic high arsenic exposure in northern Argentina." *Environ Health Perspect* 106 (6):355-9.

Cunnington, A. J. 2001. "What's so bad about teenage pregnancy?" *J Fam Plann Reprod Health Care* 27 (1):36-41.

Dietz, P. M., W. M. Callaghan, M. E. Cogswell, B. Morrow, C. Ferre, and L. A. Schieve. 2006. "Combined effects of prepregnancy body mass index and weight gain during pregnancy on the risk of preterm delivery." *Epidemiology* 17 (2):170-7. doi: 10.1097/01.ede.0000198470.26932.9a.

Fraser, A. M., J. E. Brockert, and R. H. Ward. 1995. "Association of young maternal age with adverse reproductive outcomes." *N Engl J Med* 332 (17):1113-7. doi: 10.1056/NEJM199504273321701.

Fry, R. C., P. Navasumrit, C. Valiathan, J. P. Svensson, B. J. Hogan, M. Luo, S. Bhattacharya, K. Kandjanapa, S. Soontararuks, S. Nookabkaew, C. Mahidol, M. Ruchirawat, and L. D. Samson. 2007. "Activation of inflammation/NF-kappaB signaling in infants born to arsenic-exposed mothers." *PLoS Genet* 3 (11):e207. doi: 10.1371/journal.pgen.0030207.

Gibbs, C. M., A. Wendt, S. Peters, and C. J. Hogue. 2012. "The impact of early age at first childbirth on maternal and infant health." *Paediatr Perinat Epidemiol* 26 Suppl 1:259-84. doi: 10.1111/j.1365-3016.2012.01290.x.

Goldenberg, R. L., J. F. Culhane, J. D. Iams, and R. Romero. 2008. "Epidemiology and causes of preterm birth." *Lancet* 371 (9606):75-84. doi: 10.1016/S0140-6736(08)60074-4.

Han, Z., O. Lutsiv, S. Mulla, A. Rosen, J. Beyene, S. D. McDonald, and Group Knowledge Synthesis. 2011. "Low gestational weight gain and the risk of preterm birth and low birthweight: a systematic review and meta-analyses." *Acta Obstet Gynecol Scand* 90 (9):935-54. doi: 10.1111/j.1600-0412.2011.01185.x.

He, W., R. J. Greenwell, D. M. Brooks, L. Calderon-Garciduenas, H. D. Beall, and J. D. Coffin. 2007. "Arsenic exposure in pregnant mice disrupts placental vasculogenesis and causes spontaneous abortion." *Toxicol Sci* 99 (1):244-53. doi: 10.1093/toxsci/kfm162.

Hediger, M. L., T. O. Scholl, D. H. Belsky, I. G. Ances, and R. W. Salmon. 1989. "Patterns of weight gain in adolescent pregnancy: effects on birth weight and preterm delivery." *Obstet Gynecol* 74 (1):6-12.

Hellman, L. M., M. Kobayashi, L. Fillisti, M. Lavenhar, and E. Cromb. 1969. "Growth and development of the human fetus prior to the twentieth week of gestation." *Am J Obstet Gynecol* 103 (6):789-800.

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

Hopenhayn-Rich, C., S. R. Browning, I. Hertz-Picciotto, C. Ferreccio, C. Peralta, and H. Gibb. 2000. "Chronic arsenic exposure and risk of infant mortality in two areas of Chile." *Environ Health Perspect* 108 (7):667-73.

Huyck, K. L., M. L. Kile, G. Mahiuddin, Q. Quamruzzaman, M. Rahman, C. V. Breton, C. B. Dobson, J. Frelich, E. Hoffman, J. Yousuf, S. Afroz, S. Islam, and D. C. Christiani. 2007.

"Maternal arsenic exposure associated with low birth weight in Bangladesh." *J Occup Environ Med* 49 (10):1097-104. doi: 10.1097/JOM.0b013e3181566ba0.

IOM, Institute of Medicine. 1990. "Nutrition during pregnancy. Part I. Weight gain." In *Nutrition during pregnancy. Part I. Weight gain*, edited by National Academy Press. Washington, DC.

Jekel, J. F., J. T. Harrison, D. R. Bancroft, N. C. Tyler, and L. V. Klerman. 1975. "A comparison of the health of index and subsequent babies born to school age mothers." *Am J Public Health* 65 (4):370-4.

Kile, M. L., A. Cardenas, E. Rodrigues, M. Mazumdar, C. Dobson, M. Golam, Q. Quamruzzaman, M. Rahman, and D. C. Christiani. 2016. "Estimating Effects of Arsenic Exposure During Pregnancy on Perinatal Outcomes in a Bangladeshi Cohort." *Epidemiology* 27 (2):173-81. doi: 10.1097/EDE.0000000000000416.

Kile, M. L., E. G. Rodrigues, M. Mazumdar, C. B. Dobson, N. Diao, M. Golam, Q. Quamruzzaman, M. Rahman, and D. C. Christiani. 2014. "A prospective cohort study of the association between drinking water arsenic exposure and self-reported maternal health symptoms during pregnancy in Bangladesh." *Environ Health* 13 (1):29. doi: 10.1186/1476-069X-13-29.

King, J. C. 2003. "The risk of maternal nutritional depletion and poor outcomes increases in early or closely spaced pregnancies." *J Nutr* 133 (5 Suppl 2):1732S-1736S.

Kinniburgh, D G, and P L Smedley. 2001. Arsenic contamination of groundwater in Bangladesh. UK: British Geological Survey.

Kramer, K. L., and J. B. Lancaster. 2010. "Teen motherhood in cross-cultural perspective." *Ann Hum Biol* 37 (5):613-28. doi: 10.3109/03014460903563434.

Lenders, C. M., T. F. McElrath, and T. O. Scholl. 2000. "Nutrition in adolescent pregnancy." *Curr Opin Pediatr* 12 (3):291-6.

Marsoosi, V., A. Jamal, and L. Eslamian. 2004. "Pre-pregnancy weight, low pregnancy weight gain, and preterm delivery." *Int J Gynaecol Obstet* 87 (1):36-7. doi: 10.1016/j.ijgo.2004.06.002.

Milton, A. H., W. Smith, B. Rahman, Z. Hasan, U. Kulsum, K. Dear, M. Rakibuddin, and A. Ali. 2005. "Chronic arsenic exposure and adverse pregnancy outcomes in bangladesh." *Epidemiology* 16 (1):82-6.

Mukherjee, S. C., K. C. Saha, S. Pati, R. N. Dutta, M. M. Rahman, M. K. Sengupta, S. Ahamed, D. Lodh, B. Das, M. A. Hossain, B. Nayak, A. Mukherjee, D. Chakraborti, S. K. Dulta, S. K. Palit, I. Kaies, A. K. Barua, and K. A. Asad. 2005. "Murshidabad--one of the nine groundwater arsenic-affected districts of West Bengal, India. Part II: dermatological, neurological, and obstetric findings." *Clin Toxicol (Phila)* 43 (7):835-48.

Murphy, V. E., R. Smith, W. B. Giles, and V. L. Clifton. 2006. "Endocrine regulation of human fetal growth: the role of the mother, placenta, and fetus." *Endocr Rev* 27 (2):141-69. doi: 10.1210/er.2005-0011.

Mwaniki, M. K., M. Atieno, J. E. Lawn, and C. R. Newton. 2012. "Long-term neurodevelopmental outcomes after intrauterine and neonatal insults: a systematic review." *Lancet* 379 (9814):445-52. doi: 10.1016/S0140-6736(11)61577-8.

Myers, S. L., D. T. Lobdell, Z. Liu, Y. Xia, H. Ren, Y. Li, R. K. Kwok, J. L. Mumford, and P. Mendola. 2010. "Maternal drinking water arsenic exposure and perinatal outcomes in inner Mongolia, China." *J Epidemiol Community Health* 64 (4):325-9. doi: 10.1136/jech.2008.084392.

Oehlert, Gary W. 1992. "A Note on the Delta Method." *The American Statistician* 46 (1):27-29. doi: 10.1080/00031305.1992.10475842.

Rager, J. E., K. A. Bailey, L. Smeester, S. K. Miller, J. S. Parker, J. E. Laine, Z. Drobna, J. Currier, C. Douillet, A. F. Olshan, M. Rubio-Andrade, M. Styblo, G. Garcia-Vargas, and R. C. Fry. 2013. "Prenatal arsenic exposure and the epigenome: Altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood." *Environ Mol Mutagen*. doi: 10.1002/em.21842.

Rahman, A., M. Vahter, E. C. Ekstrom, M. Rahman, A. H. Golam Mustafa, M. A. Wahed, M. Yunus, and L. A. Persson. 2007. "Association of arsenic exposure during pregnancy with fetal loss and infant death: a cohort study in Bangladesh." *Am J Epidemiol* 165 (12):1389-96. doi: 10.1093/aje/kwm025.

Raj, A., and U. Boehmer. 2013. "Girl child marriage and its association with national rates of HIV, maternal health, and infant mortality across 97 countries." *Violence Against Women* 19 (4):536-51. doi: 10.1177/1077801213487747.

Raj, A., N. Saggurti, D. Balaiah, and J. G. Silverman. 2009. "Prevalence of child marriage and its effect on fertility and fertility-control outcomes of young women in India: a cross-sectional, observational study." *Lancet* 373 (9678):1883-9. doi: 10.1016/S0140-6736(09)60246-4.

Raj, A., N. Saggurti, M. Winter, A. Labonte, M. R. Decker, D. Balaiah, and J. G. Silverman. 2010. "The effect of maternal child marriage on morbidity and mortality of children under 5 in India: cross sectional study of a nationally representative sample." *BMJ* 340:b4258. doi: 10.1136/bmj.b4258.

Rasmussen, K. M., and A. L. Yaktine. 2009. "Weight Gain During Pregnancy: Reexamining the Guidelines." In *Weight Gain During Pregnancy: Reexamining the Guidelines*, edited by K. M. Rasmussen and A. L. Yaktine. Washington (DC).

Robinson, H. P., and J. E. Fleming. 1975. "A critical evaluation of sonar "crown-rump length" measurements." *Br J Obstet Gynaecol* 82 (9):702-10.

Rudge, C. V., H. B. Rollin, C. M. Nogueira, Y. Thomassen, M. C. Rudge, and J. O. Odland. 2009. "The placenta as a barrier for toxic and essential elements in paired maternal and cord blood samples of South African delivering women." *J Environ Monit* 11 (7):1322-30. doi: 10.1039/b903805a.

Sayem, A. M., and A. T. Nury. 2011. "Factors associated with teenage marital pregnancy among Bangladeshi women." *Reprod Health* 8:16. doi: 10.1186/1742-4755-8-16.

Schieve, L. A., M. E. Cogswell, K. S. Scanlon, G. Perry, C. Ferre, C. Blackmore-Prince, S. M. Yu, and D. Rosenberg. 2000. "Prepregnancy body mass index and pregnancy weight gain: associations with preterm delivery. The NMIHS Collaborative Study Group." *Obstet Gynecol* 96 (2):194-200.

Scholl, T. O., and M. L. Hediger. 1993. "A review of the epidemiology of nutrition and adolescent pregnancy: maternal growth during pregnancy and its effect on the fetus." *J Am Coll Nutr* 12 (2):101-7.

Scholl, T. O., and M. L. Hediger. 1995. "Weight gain, nutrition, and pregnancy outcome: findings from the Camden study of teenage and minority gravidas." *Semin Perinatol* 19 (3):171-81.

Scholl, T. O., M. L. Hediger, and D. H. Belsky. 1994. "Prenatal care and maternal health during adolescent pregnancy: a review and meta-analysis." *J Adolesc Health* 15 (6):444-56.

Shah, R., L. C. Mullany, G. L. Darmstadt, I. Mannan, S. M. Rahman, R. R. Talukder, J. A. Applegate, N. Begum, D. Mitra, S. E. Arifeen, A. H. Baqui, and AHNMo Study Group in Bangladesh Proj. 2014. "Incidence and risk factors of preterm birth in a rural Bangladeshi cohort." *BMC Pediatr* 14:112. doi: 10.1186/1471-2431-14-112.

Shawky, S., and W. Milaat. 2000. "Early teenage marriage and subsequent pregnancy outcome." *East Mediterr Health J* 6 (1):46-54.

Shi, X., J. D. Ayotte, A. Onda, S. Miller, J. Rees, D. Gilbert-Diamond, T. Onega, J. Gui, M. Karagas, and J. Moeschler. 2015. "Geospatial association between adverse birth outcomes and arsenic in groundwater in New Hampshire, USA." *Environ Geochem Health* 37 (2):333-51. doi: 10.1007/s10653-014-9651-2.

Siega-Riz, A. M., L. S. Adair, and C. J. Hobel. 1994. "Institute of Medicine maternal weight gain recommendations and pregnancy outcome in a predominantly Hispanic population." *Obstet Gynecol* 84 (4):565-73.

Siega-Riz, A. M., L. S. Adair, and C. J. Hobel. 1996. "Maternal underweight status and inadequate rate of weight gain during the third trimester of pregnancy increases the risk of preterm delivery." *J Nutr* 126 (1):146-53.

Slotnick, M. J., J. R. Meliker, and J. O. Nriagu. 2006. "Effects of time and point-of-use devices on arsenic levels in Southeastern Michigan drinking water, USA." *Sci Total Environ* 369 (1-3):42-50. doi: 10.1016/j.scitotenv.2006.04.021.

Sohel, N., P. S. Kanaroglou, L. A. Persson, M. Z. Haq, M. Rahman, and M. Vahter. 2010. "Spatial modelling of individual arsenic exposure via well water: evaluation of arsenic in urine, main water source and influence of neighbourhood water sources in rural Bangladesh." *J Environ Monit* 12 (6):1341-8. doi: 10.1039/c001708f.

Spinillo, A., E. Capuzzo, G. Piazzi, A. Ferrari, V. Morales, and M. Di Mario. 1998. "Risk for spontaneous preterm delivery by combined body mass index and gestational weight gain patterns." *Acta Obstet Gynecol Scand* 77 (1):32-6.

UNFPA, (United Nations Population Fund). 2012. Marrying Too Young, End Child Marriage. New York.

UNFPA, (United Nations Population Fund). 2013. State of the World's Children, 2013.

Vall, O., M. Gomez-Culebras, O. Garcia-Algar, X. Joya, D. Velez, E. Rodriguez-Carrasco, and C. Puig. 2012. "Assessment of prenatal exposure to arsenic in Tenerife Island." *PLoS One* 7 (11):e50463. doi: 10.1371/journal.pone.0050463.

Varma, T. R. 1984. "Maternal weight and weight gain in pregnancy and obstetric outcome." *Int J Gynaecol Obstet* 22 (2):161-6.

von Ehrenstein, O. S., D. N. Guha Mazumder, M. Hira-Smith, N. Ghosh, Y. Yuan, G. Windham, A. Ghosh, R. Haque, S. Lahiri, D. Kalman, S. Das, and A. H. Smith. 2006. "Pregnancy outcomes, infant mortality, and arsenic in drinking water in West Bengal, India." *Am J Epidemiol* 163 (7):662-9. doi: 10.1093/aje/kwj089.

Wallace, J. M., R. P. Aitken, J. S. Milne, and W. W. Hay, Jr. 2004. "Nutritionally mediated placental growth restriction in the growing adolescent: consequences for the fetus." *Biol Reprod* 71 (4):1055-62. doi: 10.1095/biolreprod.104.030965.

Xu, L., K. Yokoyama, Y. Tian, F. Y. Piao, F. Kitamura, H. Kida, and P. Wang. 2011. "Decrease in birth weight and gestational age by arsenic among the newborn in Shanghai, China." *Nihon Koshu Eisei Zasshi* 58 (2):89-95.

Yang, C. Y., C. C. Chang, S. S. Tsai, H. Y. Chuang, C. K. Ho, and T. N. Wu. 2003. "Arsenic in drinking water and adverse pregnancy outcome in an arseniasis-endemic area in northeastern Taiwan." *Environ Res* 91 (1):29-34.

Zou, G. 2004. "A modified poisson regression approach to prospective studies with binary data." *Am J Epidemiol* 159 (7):702-6.

2. Chapter 2: Investigating causal relation between prenatal arsenic exposure and birthweight: Are smaller infants more susceptible?

Authors: Mohammad Rahman[1], Linda Valeri[2], Molly L. Kile[3], Maitreyi Mazumdar[1], Golam Mostofa[4], Qazi Qamruzzaman[4], Mahmudur Rahman[4], Andrea Baccarelli[5], Liming Liang[6], Russ Hauser[1,7], David C. Christiani[1,7]

Author Affiliations:

[1]Harvard T.H. Chan School of Public Health, Department of Environmental Health, Boston, Massachusetts, USA

[2]McLean Hospital, Belmont, Massachusetts, USA and Harvard Medical School, Boston, Massachusetts, USA

[3] Oregon State University, College of Public Health and Human Sciences, Corvallis, Oregon, USA

[4]Dhaka Community Hospital Trust, Dhaka, Bangladesh

[5]Columbia University, Mailman School of Public Health, Department of Environmental Health, New York, NY, USA

[6]Harvard T.H. Chan School of Public Health, Department of Biostatistics, Boston, Massachusetts, USA

[7]Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, Massachusetts, USA

Abstract

Background: Shortened gestation and intrauterine growth restriction (IUGR) are the two main processes that govern birthweight. Research is limited in understanding the causal relation between prenatal arsenic exposure and birthweight, and whether the magnitude of this relation varies across birthweight percentiles.

Objectives: Using quantile causal mediation analysis approach, we aimed to determine the relation between prenatal arsenic exposure and birthweight in relation to shortened gestation and IUGR, and to identify infants who might be disproportionately affected by arsenic exposure.

Methods: In a longitudinal Bangladeshi birth cohort, we measured arsenic in drinking water (n=1,181) collected at enrollment and maternal toenails (n=1,104) collected within one-month postpartum using inductively coupled plasma mass spectrometry, determined gestational age using ultrasound at ≤16 weeks' gestation, and collected demographic information using structured questionnaire.

Results: Of 1,184 singleton livebirths, 16.4% (n=194) were low birthweight (<2500 g), 21.9% (n=259) preterm (<37 weeks gestation) and 9.2% (n=109) both low birthweight and preterm. Prenatal arsenic exposure was associated with decreased birthweight, and the magnitude of the associations varied across birthweight percentiles. While infants of all birth sizes experienced decreased birthweight because of shortening of gestational age (β range: 5th percentile= −20.5g [95% CI: −30.7, −10.9] to 95th percentile= −8.9g [95% CI: −17.2, −1.9] per natural log water arsenic increase), the smaller infants experienced decreased birthweight because of both shortening of gestational age and IUGR (β range: 5th percentile= −35.2g [95% CI: −62.2, −7.6] to 25th percentile= −15.0g [95% CI: −31.9, −0.4] per natural log water arsenic increase). Similar associations were observed for maternal toenail arsenic.

Conclusions: The susceptibility of prenatal arsenic exposure varied by infant birth sizes, placing smaller infants at greater risk of decreased birthweight due to both shortened gestation and IUGR. It is important to mitigate prenatal arsenic exposure to improve perinatal outcomes in Bangladesh.

Introduction

Low birthweight, weighing <2500 g at birth, is an important population indicator for neonatal mortality and a determinant of infant and childhood morbidity.(Kramer 1987) Each year, an estimated 21 million infants are born with low birthweight worldwide, more than half of them are in South Asia.(WHO and UNICEF 2004) In Bangladesh, the incidence of low birthweight is estimated to be 22%, which is among the highest in the world(Lee et al. 2013), while in rural areas the estimates are as high as 31−47%.(Hosain et al. 2006) Furthermore, about 40 million people in Bangladesh, or a quarter of the country's population, are still exposed to arsenic through drinking water sourced from groundwater above the World Health Organization (WHO) recommended level of <10 μg/L.(Loewenberg 2016) Arsenic exposure is much more highly prevalent in rural areas, where 97% of the population relies on groundwater for drinking purposes.(Flanagan, Johnston, and Zheng 2012)

Arsenic can cross the placenta readily,(Concha, Nermell, and Vahter 1998) and several studies have reported that prenatal arsenic exposure is associated with spontaneous abortion,(Rahman et al. 2010, Milton et al. 2005, He et al. 2007), preterm birth,(Ahmad et al. 2001, Chakraborti et al. 2003, Yang et al. 2003) and intrauterine growth restriction.(Thomas et al. 2015, Kippler et al. 2012, Llanos and Ronco 2009, Claus Henn et al. 2016) But the evidence for the association between arsenic and birthweight is less consistent. While ten studies (Ahamed et al. 2006, Chakraborti et al. 2004, Gelmann et al. 2013, Kwok, Kaufmann, and Jakariya 2006,

Mukherjee et al. 2005, Myers et al. 2010, Shirai et al. 2010, Vall et al. 2012, Hu et al. 2015) reported a null or positive associations between prenatal arsenic exposure and birthweight, 14 other studies (Bloom et al. 2016, Chakraborti et al. 2003, Guan et al. 2012, Hopenhayn et al. 2003, Huyck et al. 2007, Kile et al. 2016, Rahman et al. 2009, Xu et al. 2011, Yang et al. 2003, Claus Henn et al. 2016, Fei et al. 2013, Gilbert-Diamond et al. 2016, McDermott et al. 2014, Llanos and Ronco 2009, Laine et al. 2015), including 6 large prospective cohort studies, reported negative associations. These negative associations were consistent, despite the use of different exposure measures in drinking water,(Hopenhayn et al. 2003, Kile et al. 2016, Yang et al. 2003) maternal urine,(Fei et al. 2013, Gilbert-Diamond et al. 2016, Rahman et al. 2009, Laine et al. 2015) toenail,(Kile et al. 2016) hair,(Huyck et al. 2007) whole blood,(Claus Henn et al. 2016, Xu et al. 2011, Guan et al. 2012), soil around home,(McDermott et al. 2014) and placental tissue,(Llanos and Ronco 2009) as well as different levels of exposure experienced by the study populations. And while these studies were largely conducted among populations with frequent exposure to higher levels of arsenic through drinking water (i.e. >10μg/L) in Bangladesh,(Huyck et al. 2007, Kile et al. 2016, Rahman et al. 2009) India,(Chakraborti et al. 2003) Chile,(Hopenhayn et al. 2003) and Taiwan,(Yang et al. 2003) several recent studies in the United States(Fei et al. 2013, Gilbert-Diamond et al. 2016, Claus Henn et al. 2016) and China(Guan et al. 2012, Xu et al. 2011) have corroborated negative associations between arsenic and birthweight among populations exposed to relatively lower levels of exposure.

Low birthweight is often caused by either shortened gestation or intrauterine growth restriction (IUGR), which is commonly assessed by small for gestational age ($<10^{th}$ percentile of the birthweight-for-gestational age distribution). Without distinguishing these two factors, it is difficult to identify the true causal determinants of low birthweight and to develop effective

public health interventions.(Kline 1989, Kramer 1987, Lee et al. 2013) And most studies have focused on birthweight as a single entity and have not used a causal framework that includes both shortened gestation and IUGR for analysis. Mediation analysis can help identify arsenic-birthweight associations in relation to both shortened gestation and IUGR, by decomposing the total effect of arsenic exposure on birthweight into indirect effect via pathways mediated through gestational age and direct effect via pathways independent of gestational age, respectively.(Valeri and Vanderweele 2013) For instance, using structural equation modeling technique, our group recently identified a negative association between prenatal arsenic exposure and decreased birthweight, which was mediated primarily via shortening of gestational age, while the effects of arsenic exposure independent of gestational age were in the positive direction and not statistically significant.(Kile et al. 2016)

Infants at the tails of birthweight distribution often suffer a disproportionate burden of perinatal mortality and morbidity.(Barker et al. 2002) If arsenic exposure disproportionately affects infants at the tails of birthweight distribution, identifying those infants could have important implications in reducing perinatal risks in this population. We hypothesized that the pathway-specific effects of arsenic exposure on birthweight would vary across birthweight percentiles, depending on whether the pathways are via gestational age or independent of gestational age. However, all previous studies have assumed a homogenous association between arsenic exposure and birthweight across birthweight distribution and summarized the effect estimates that might have differed across the range of birthweight percentiles, including those with opposing signs. Traditional modeling approaches cannot capture our question, but causal mediation modeling techniques can be combined with quantile regression and would allow us to

identify pathway-specific effects of arsenic exposure on birthweight across birthweight percentiles.(Imai, Keele, and Tingley 2010)

The aim of this paper was to determine the causal association between prenatal arsenic exposure and birthweight in relation to shortened gestation and intrauterine growth restriction in order to identify infants who might be disproportionately affected by arsenic exposure in a longitudinal Bangladeshi birth cohort.

Methods

Study population

A longitudinal birth cohort was established in Bangladesh between 2008-2011. The details of this study, including recruitment and enrollment process were previously described.(Kile et al. 2014) Briefly, women were eligible to participate if they were 18 years or older with an ultrasound confirmed singleton pregnancy of ≤16 weeks' gestation, used a tube well as their primary source of drinking water and had been using the same drinking water source for more than six months, and intended to live in her current residence throughout her pregnancy. Of 1,613 pregnant women innitially recruited, our analysis included data from 1,181 singleton livebirths between 28−42 weeks' gestation after exclusions due to loss of contact (n=99), participation withdrawn (n=121), miscarriage (n=132), stillbirth (n=72), twin births (n=5), and missing data on water arsenic (n=2) and secondhand smoking (n=2). All subjects provided written informed consent before participation. Participants were informed and counseled on safe drinking water options if their water samples contained arsenic above Bangladeshi standard (i.e.<50μg/L). Prenatal care and multivitamins were provided to all participants. All protocols were reviewed and approved by the Human Research Committees at Harvard T.H. Chan School of Public Health and Dhaka Community Hospital Trust.

Exposure Assessment

Arsenic was measured in drinking water (n=1,181) from tubewells that women identified as their principal water source at the time of enrollment. Details of sample collection and measurement procedures have been previously described.(Kile et al. 2016) Briefly, water samples were collected in a 50-ml polypropylene tubes (BD Falcon, BD Bioscience, Bedford,MA) and preserved with Reagent Grade nitric acid (Merck, Germany) to a pH<2. Samples were kept at room temperature until analysis by inductively coupled plasma mass spectrometry (ICP-MS) following US EPA method 200.8 (Environmental Laboratory Services, North Syracuse, New York). The average percent recovery of arsenic was 101% (range:92%−110%). Samples with arsenic concentrations below the limit of detection (LOD) (n=242) ranging from 0.5-1.0 μg/L were reassigned half the value of the LOD for statistical analysis.

Arsenic was also measured in maternal toenails collected ≤1-month post-partum. Toenail samples were sonicated in 1% Triton X-100 solution (Sigma-Aldrich, Inc., St. Louis,MO) and rinsed repeatedly with Milli-Q water (Millipore Corporation, Billerica,MA) to remove external contamination before microwave acid digestion using Trace Select Ultra Pure nitric acid (Fischer Scientific, Pittsburgh,PA). Digested samples were diluted with Milli-Q water before analyzing for total arsenic using ICP-MS. The reported arsenic concentrations were blank-corrected and normalized using arsenic concentration of certified human hair reference material (CRM Hair; Shanghi Institute of Nuclear Research, China). Toenail clippings were collected from 1,155 samples, 1,104 of them were used in the analysis after excluding samples with low mass (≤5mg) and/or high relative standard deviation (≥25%). Samples with arsenic concentrations below the LOD ranging from 0.09−0.7ng/L were reassigned half the value of LOD for statistical analysis.

Outcome and covariates

The study involved four scheduled visits occurring at the time of enrollment, around 28th weeks' gestation, at the time of delivery, and ≤1 month post-partum. During those visits trained interviewers used structured questionnaires to collect demographic, medical, and environmental information. Birthweight was measured within 120 minutes after delivery on a pediatric scale calibrated and rounded to the nearest 10 grams before each measurement. All pregnancies were dated either by crown-rump length (CRL) at 7−16 weeks or mean sac diameter (MSD) at 4−6 weeks. Fetal CRL and MSD were measured by a trained family physician by ultrasound using the formulae proposed by Robinson and Hellman, respectively.(Hellman et al. 1969, Robinson and Fleming 1975) Maternal blood samples were collected at the time of enrollment to measure hemoglobin levels. Maternal height and weight were recorded at first clinic visit and at monthly house visits following enrollment, when they also received prenatal vitamins.

Statistical Analysis

Arsenic concentrations in drinking water, and maternal toenail were right skewed and subsequently transformed to their natural log. Mean birthweight across categories of all covariates were analyzed using T-test or analysis variance (ANOVA) in bivariate models. The distribution of birthweight was checked and a histogram indicated no gross violation for normality assumption.

Mediation analysis was implemented considering maternal hemoglobin level at the time of enrollment and gestational age as mediators following the method described by VanderWeele and Vansteelandt.(VanderWeele and Vansteelandt 2014) The directed acyclic graph (DAG) in Figure 1(a-e) explains the conceptual model for causal mediation analysis. We hypothesized that maternal hemoglobin at enrollment and gestational age lie within the causal pathways between arsenic exposure and birthweight with maternal hemoglobin level temporally preceding

gestational age. We modeled maternal hemoglobin and gestational age using linear regression (mediator models). Models of hemoglobin were adjusted for arsenic, maternal age, education, number of past pregnancies, enrollment BMI, secondhand smoke, and infant sex, while models of gestational age were additionally adjusted for maternal hemoglobin. We used quantile regressions to model percentiles of birthweight (outcome models) adjusting for arsenic, hemoglobin, maternal age, education, number of past pregnancies, enrollment BMI, secondhand smoke, infant sex, arsenic-hemoglobin interaction term, and gestational age (for joint mediation effect or mediation effect through gestational age only). The outcome and mediator models we combined to estimate the direct, indirect and total effects of arsenic on birthweight at 5^{th}-95^{th} percentiles of birthweight distribution. Analyses were repeated for drinking water arsenic and

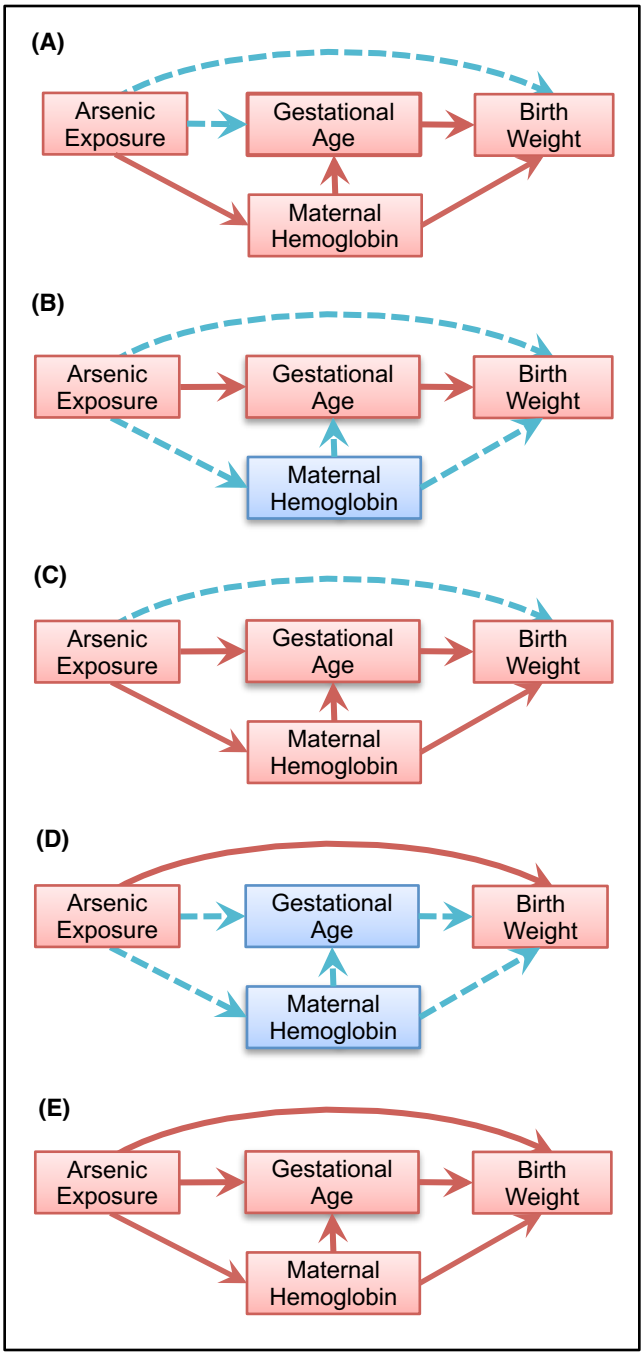

Figure 2.1. Conceptual model for causal mediation analysis showing associations between in-utero arsenic exposure and birthweight. Pathways highlighted in red with solid lines indicate: A) indirect effect mediated via hemoglobin, B) indirect effect mediated via gestational age, C) total indirect effects via both hemoglobin and gestational age, D) Direct effect via pathways other than hemoglobin and gestational age, and E) total effects

maternal toenail arsenic. Direct and indirect effects were averaged across all individuals. Bias corrected standard errors were computed from 1000 bootstrap replications.

Analyses assume that conditional on the covariates, there is no confounding of 1) the exposure-outcome relation, 2) exposure-mediator relation, (3) mediator-outcome relation and that (4) there is no effect of the exposure that itself confounds the mediator-outcome relation.(Valeri and Vanderweele 2013) While assumption 4 may appear not to hold for a single mediator, it holds jointly for both mediators. Analyses were implemented with R 3.2.3 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Table 1 summarizes demographic characteristics of the study population and how these characteristics are associated with birthweight. Average birthweight was 2837g (standard deviation: 408g; range: 800-4,800g). Nearly 22% (n=259) of the infants were born preterm (<37 weeks of gestation), 19% (n=223) small-for-gestational age,(Villar et al. 2014) 16.4% (n=194) low birthweight and 9.2% (n=109) both preterm and low birthweight. Arsenic exposures were relatively modest but spanned a wide range. The median concentrations of arsenic exposure were 2.2μg/L (range: <LOD−1400μg/L) for drinking water arsenic, and 1.2μg/g (range: <LOD−46.6μg/g) for maternal toenail arsenic. Drinking water arsenic concentration showed a modest correlation with maternal toenail arsenic ($\sigma_{spearman}$=0.50).

The adjusted indirect, direct, and total effects of drinking water arsenic exposure on birth weight across birthweight percentiles were presented in Table 2 and Figure 2, whereas the adjusted indirect, direct, and total effects of toenail arsenic on birth weight across birthweight percentiles were presented in table 3 and Figure 3. Our results suggested a heterogeneous relation between arsenic and birthweight across birthweight quantiles, which involved pathways

Table 2.1. Distribution of selected characteristics in the study cohort

Characteristic	N (%)	**Birthweight** (g) Mean (SD)[a]
Age group (years)		
18−20	470 (39.8)	2836 (383)
21−25	438 (37.1)	2842 (428)
26−41	273 (23.1)	2831 (420)
Gestational Age		
28−<37 weeks	259 (21.9)	2566 (498)
37−42 weeks	922 (78.0)	2913 (343)
Enrollment BMI		
≤18·5 kg/m^2	335 (28.4)	2775 (419)
>18·5 to ≤25.0 kg/m^2	738 (62.5)	2838 (392)
>25.0 kg/m^2	108 (9.1)	3013 (431)
Birth Type		
Normal	766 (65.9)	2771 (417)
Cesarean	415 (35.1)	2959 (361)
Infant Sex		
Female	598 (50.6)	2873 (394)
Male	583 (49.4)	2799 (420)
No. of past pregnancies		
0	475 (40.2)	2859 (389)
1	353 (29.9)	2855 (418)
≥2	353 (29.9)	2790 (422)
Maternal education, n (%)		
No formal education	172 (14.6)	2712 (402)
Primary education	380 (32.2)	2844 (403)
Secondary or higher	629 (53.2)	2866 (407)
Secondhand smoke, n (%)		
Yes	495 (41.9)	2794 (405)
No	685 (56.1)	2868 (409)
Blood hemoglobin (gm/L)		
7.9−10.9	528 (44.7)	2828 (441)
11−16.3	653 (55.3)	2844 (380)
Drinking water arsenic (μg/L)		
<LOD−10	727 (61.6)	2875 (351)
>10−50	194 (16.4)	2750 (493)
>50−1400	259 (22.0)	2796 (472)
Toenail arsenic (μg/g)		
<LOD−1.2	554 (50.3)	2861 (357)
>1.2−2.3	217 (19.7)	2792 (448)
>2.3−46.5	331 (30.0)	2854 (431)

Table 2.2. Indirect, Direct, and Total Effects of Natural Log-transformed Arsenic Concentrations in Drinking Water (μg/L) on Quantiles of Birthweight (g) Considering Maternal Hemoglobin Level at Enrollment (g/dL) and Gestational Age (weeks) as Mediators (n=1,181)

Percentile	**Birth Weight** (g)	**Indirect Effects via Hemoglobin*** (95% CI)	**Indirect Effects via Gestational Age*** (95% CI)	**Joint Indirect Effects** (95% CI)	**Direct Effects** (95% CI)	**Total Effects** (95% CI)
0.05	2100	-3.1 (-10.6, 5.4)	-20.5 (-30.7, -10.9)	-23.6 (-33.8, -13.9)	-35.2 (-62.2, -7.6)	-58.8 (-86.6, -31.7)
0.10	2300	-0.6 (-6.7, 5.5)	-20.1 (-28.7, -12.2)	-20.7 (-28.8, -12.7)	-31.7 (-56.4, -9.0)	-52.4 (-77.4, -31.2)
0.15	2450	0.7 (-4.6, 5.6)	-21.1 (-29, -13.9)	-20.4 (-28.4, -13.2)	-24.6 (-44.2, -6.1)	-45.0 (-64.1, -27.4)
0.20	2520	0.9 (-2.9, 4.8)	-21.2 (-28.7, -14.2)	-20.4 (-27.7, -13.2)	-19.8 (-40.0, -1.9)	-40.2 (-59.7, -23.0)
0.25	2600	0.4 (-2.8, 3.7)	-20.0 (-26.8, -13.7)	-19.6 (-26.8, -13.2)	-15.0 (-31.9, -0.4)	-34.6 (-52.2, -19.9)
0.30	2710	0.4 (-2.3, 3.2)	-19.1 (-25.8, -13.1)	-18.7 (-25.3, -12.7)	-11.4 (-24.7, 0.7)	-30.1 (-45.0, -17.4)
0.35	2760	0.4 (-2.0, 3.0)	-18.0 (-24.2, -11.8)	-17.6 (-24, -11.4)	-9.5 (-24.6, 4.9)	-27.1 (-42.9, -11.8)
0.40	2800	0.5 (-2.1, 3.1)	-16.6 (-23.1, -11.1)	-16.1 (-22.6, -10.1)	-6.5 (-21.4, 7.5)	-22.6 (-39.2, -8.8)
0.45	2820	0.7 (-1.9, 3.5)	-15.8 (-21.7, -10.4)	-15.1 (-21.6, -9.4)	-2.9 (-18.9, 12.8)	-18.0 (-33.8, -2.2)
0.50	2860	1.1 (-1.9, 3.7)	-15.3 (-21.5, -9.6)	-14.3 (-21.4, -8.1)	1.8 (-14.4, 17.8)	-12.4 (-30.5, 4.1)
0.55	2890	1.3 (-1.9, 4)	-14.4 (-20.9, -8.6)	-13.1 (-20.4, -6.4)	4.8 (-11.2, 20.6)	-8.3 (-25.0, 7.2)
0.60	2910	1.3 (-2.5, 4.7)	-13.1 (-19.9, -7.6)	-11.9 (-19.3, -5.4)	6.7 (-10.4, 23.2)	-5.2 (-23.2, 11.1)
0.65	3010	0.9 (-3.4, 5.2)	-11.7 (-18.5, -6.2)	-10.8 (-18.1, -4.6)	8.8 (-10.8, 25.3)	-2.0 (-23.9, 15.0)
0.70	3060	0.4 (-4.0, 5.5)	-10.7 (-16.7, -6.0)	-10.3 (-16.6, -4.4)	10.5 (-10.5, 28.5)	0.1 (-22.9, 18.4)
0.75	3100	-0.3 (-4.6, 4.6)	-9.7 (-15.0, -5.0)	-10.0 (-15.9, -4.6)	11.2 (-11.0, 31.0)	1.2 (-22.1, 21.9)
0.80	3160	-0.3 (-4.1, 4.5)	-9.3 (-14.9, -4.8)	-9.6 (-15.5, -4.4)	11.5 (-11.4, 33.8)	1.9 (-22.7, 24.3)
0.85	3200	0.9 (-3.3, 4.8)	-10.1 (-16.3, -5.0)	-9.2 (-15.2, -3.6)	15.7 (-7.7, 37.3)	6.5 (-16.9, 28.0)
0.90	3250	3.3 (-0.5, 7.1)	-11.3 (-18.2, -5.0)	-8.0 (-15.7, -0.9)	18.2 (-7.2, 39.6)	10.2 (-14.7, 32.0)
0.95	3450	4.5 (-1.0, 10.6)	-8.9 (-17.2, -1.9)	-4.4 (-14.3, 4.4)	20.9 (-9.6, 58.5)	16.5 (-12.9, 53.8)

* Models for maternal hemoglobin and gestational age (mediator models) were adjusted for natural log drinking water arsenic concentrations maternal age, education, secondhand smoking, enrollment BMI, number of past pregnancies, and infant sex

* Models for birthweight (outcome model) were additionally adjusted for maternal hemoglobin level

* Models for birthweight (outcome model) were additionally adjusted for maternal hemoglobin level and gestational age.

Direct and indirect effects were averaged over all individuals

Bias corrected standard errors were derived from 1,000 bootstraps

mediated through gestational age as well as pathways independent of gestational age. The observed causal relations were consistent regardless of whether arsenic was measured in drinking water or maternal toenails.

Additionally, our results showed that arsenic exposure had no significant indirect effect on birthweight through maternal hemoglobin level across birthweight percentiles (Figure 2A, 3A). The indirect effect mediated through gestational age, beyond hemoglobin level were negative and statistically significant for all infants irrespective of birth sizes, although the associations were stronger among smaller infants (Figure 2B, 3B). For instance, among infants with birthweight <2100g (i.e.5thpercentile), a unit increase in natural log water arsenic exposure was associated with 20.5g (95%CI: −30.7, −10.9) decrease in birthweight mediated through gestational age beyond hemoglobin level, while among

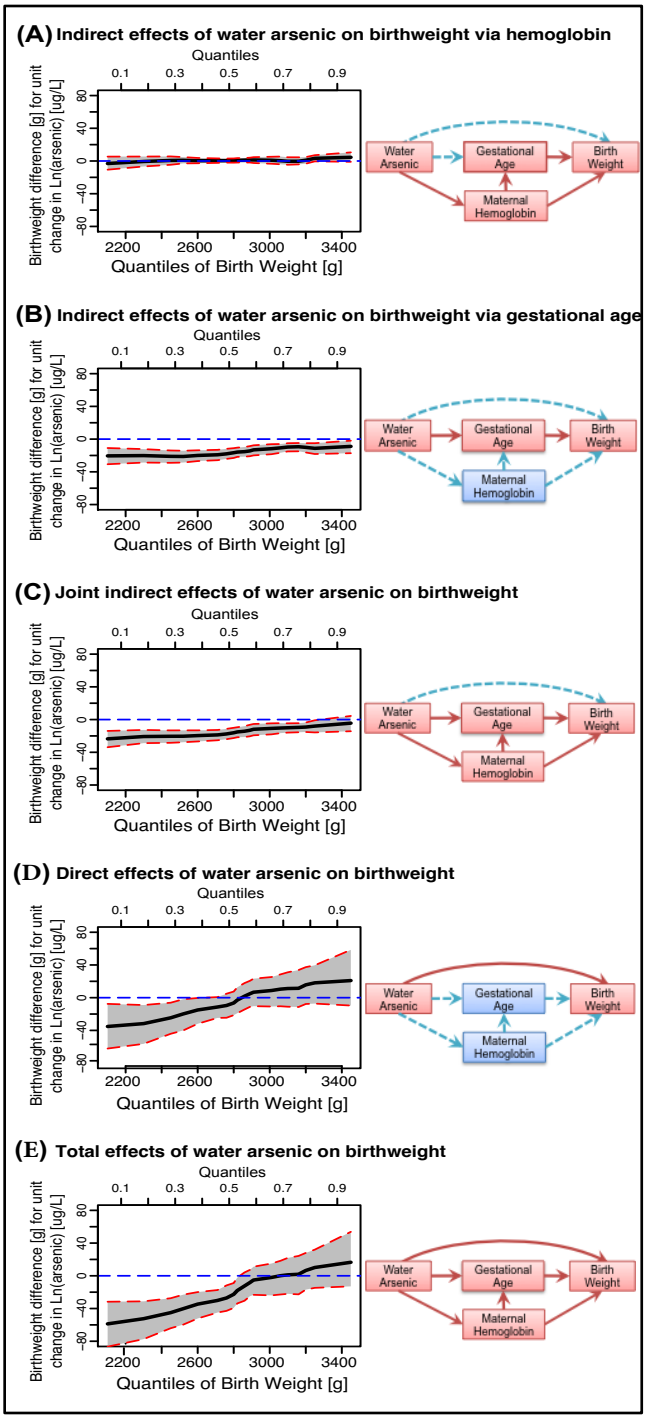

Figure 2.2. Indirect, direct and total effects of drinking water arsenic exposure on birthweight considering maternal hemoglobin and gestational age as mediators. Solid black lines surrounded by shaded areas represent effect estimates and 95% confidence intervals across birthweight quantiles. Horizontal blue dashed lines show reference values. Directed acyclic graphs illustrate corresponding active (red with solid lines) and inactive (blue with dashed lines) causal pathways.

infants born with birthweight >3453g (i.e. 95th percentile), for the same exposure the change in birthweight was −8.9g (95%CI: −17.2, −1.9) (Table 2). Similar association was observed for maternal toenail arsenic (Table 3).

Table 2.3. Indirect, Direct, and Total Effects of Natural Log-transformed Arsenic Concentrations in Maternal Toenail (μg/L) on Quantiles of Birthweight (g) Considering Maternal Hemoglobin Level at Enrollment (g/dL) and Gestational Age (weeks) as Mediators

Percentile	**Birth Weight** (g)	**Indirect Effects via Hemoglobin*** (95% CI)	**Indirect Effects via Gestational Age*** (95% CI)	**Joint Indirect Effects** (95% CI)	**Direct Effects** (95% CI)	**Total Effects** (95% CI)
0.05	2100	-2.2 (-9.9, 6.1)	-19 (-30, -9.7)	-21.1 (-31.1, -11.9)	-29.5 (-56.9, 2.3)	-50.6 (-77, -22.3)
0.10	2300	-0.5 (-6.1, 5.7)	-17.5 (-27.4, -9.1)	-17.9 (-26.7, -10.4)	-27.9 (-51.9, -5.9)	-45.8 (-68.6, -26.4)
0.15	2450	0.4 (-4.9, 5.5)	-17.6 (-26.2, -10.3)	-17.2 (-25.5, -10.2)	-21.8 (-43.8, -2.4)	-39.0 (-60.7, -20.9)
0.20	2520	0.4 (-4.0, 5.0)	-17.5 (-25.2, -10.6)	-17 (-24.8, -10.3)	-16.9 (-36, 0.9)	-34.0 (-51.7, -17.2)
0.25	2600	0.1 (-3.3, 3.8)	-16.5 (-23.6, -10.2)	-16.4 (-23.6, -10.1)	-12.8 (-29.6, 2.3)	-29.2 (-47.1, -14.0)
0.30	2710	0.1 (-2.9, 3.1)	-15.6 (-22.1, -9.9)	-15.4 (-22.5, -9.6)	-9.9 (-25.4, 4.6)	-25.3 (-41.8, -10.2)
0.35	2760	0.2 (-2.5, 3.1)	-14.3 (-20.7, -8.9)	-14.2 (-21, -8.5)	-6.7 (-22.8, 8.6)	-20.9 (-39.3, -5.2)
0.40	2800	0.3 (-2.3, 3.1)	-13.2 (-19.5, -8.0)	-12.9 (-19.3, -7.5)	-3.2 (-19.4, 12.9)	-16.0 (-33.0, 0.00)
0.45	2820	0.7 (-2.0, 3.4)	-12.7 (-18.5, -7.4)	-12 (-18.3, -6.3)	1.2 (-16.5, 17.2)	-10.8 (-29.2, 5.1)
0.50	2860	1.3 (-2.0, 4.1)	-12.2 (-17.9, -6.9)	-10.9 (-17.5, -5)	6.1 (-12.9, 23.1)	-4.8 (-23.1, 11.6)
0.55	2890	1.5 (-2.5, 4.6)	-11.2 (-17.4, -6.0)	-9.7 (-16.7, -4.2)	9.2 (-8.2, 24.7)	-0.4 (-19.6, 14.6)
0.60	2910	1.4 (-2.5, 5.0)	-10 (-15.9, -5.1)	-8.6 (-15.6, -3)	11.3 (-6.4, 27.0)	2.7 (-15.6, 18.6)
0.65	3010	0.9 (-3.5, 5.3)	-8.8 (-14.1, -4.3)	-7.9 (-14.5, -2.7)	12.9 (-6.8, 30.1)	5.0 (-13.8, 21.4)
0.70	3060	0.3 (-4.2, 5.3)	-8.1 (-13, -3.7)	-7.9 (-13.7, -2.5)	13.3 (-9, 30.9)	5.4 (-16.7, 23.9)
0.75	3100	-0.2 (-4.6, 4.9)	-7.5 (-12.4, -3.5)	-7.7 (-13.2, -3.0)	13.4 (-10.5, 33.7)	5.7 (-18.6, 26.8)
0.80	3160	-0.2 (-4.4, 4.5)	-7.2 (-12.2, -2.9)	-7.4 (-12.7, -2.4)	13.8 (-8.1, 33.3)	6.4 (-14.8, 25.6)
0.85	3200	1.1 (-3.5, 5.7)	-8.2 (-14.1, -3.4)	-7 (-12.5, -1.8)	16.3 (-6.6, 35.5)	9.3 (-13.6, 28.1)
0.90	3250	3.3 (-0.7, 7.6)	-8.8 (-15.4, -2.9)	-5.4 (-12.8, 1.4)	18.5 (-6.3, 38.8)	13.1 (-10.8, 34.1)
0.95	3450	4.4 (-2, 10.1)	-6.3 (-15, 0.4)	-1.9 (-11.4, 7.3)	22.6 (-9.9, 58.7)	20.7 (-11.2, 54.5)

* Models for maternal hemoglobin and gestational age (mediator models) were adjusted for natural log of maternal toenail arsenic, maternal age, education, secondhand smoking, enrollment BMI, number of past pregnancies, and infant sex
* Models for birthweight (outcome model) were additionally adjusted for maternal hemoglobin level
* Models for birthweight (outcome model) were additionally adjusted for maternal hemoglobin level and gestational age.
Direct and indirect effects were averaged over all individuals
Bias corrected standard errors were derived from 1,000 bootstraps

Our results indicate that via shortening of gestational age, arsenic exposure will shift the entire distribution of birthweight to the left with an extended lower tail. The joint indirect effect mediated through both hemoglobin and gestational age across birthweight quantiles were predominantly contributed by the indirect effect mediated through gestational age (Figure 2C, 3C).

The direct effect of arsenic on birthweight through pathways independent of gestational age and hemoglobin level across birthweight percentiles were heterogeneous and bidirectional (Figure 2D, 3D). For instance, via direct pathways, each unit increase in natural log water arsenic exposure was associated with 35.2g (95%CI: −62.1, −7.6) decrease in birthweight among infants with birthweight <2100g (5thpercentile), whereas the decrease was 15.0g (95%CI: −31.9, −0.4 g) among infants with birthweight around 2607g

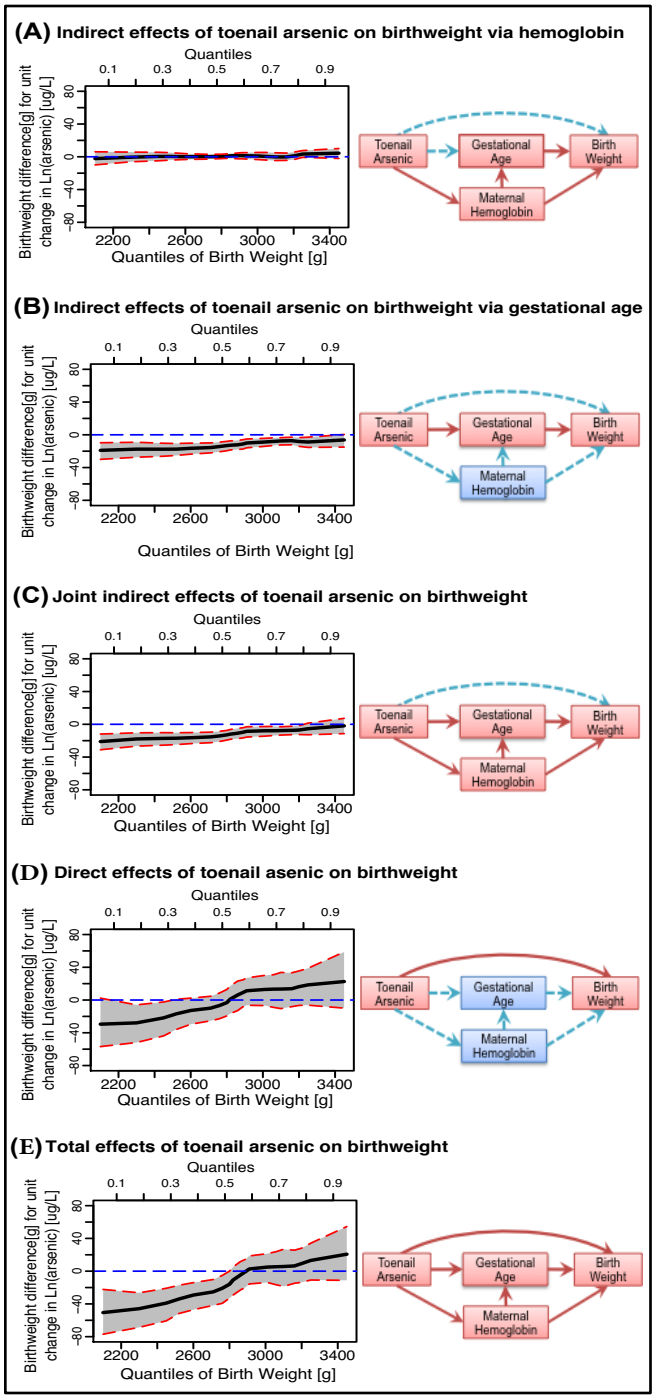

Figure 2.3. Indirect, direct and total effects of maternal toenail arsenic exposure on birthweight considering maternal hemoglobin and gestational age as mediators. Solid black lines surrounded by shaded areas represent effect estimates and 95% confidence intervals across birthweight quantiles. Horizontal blue dashed lines show reference values. Directed acyclic graphs illustrate corresponding active (red with solid lines) and inactive (blue with dashed lines) causal pathways

(25thpercentile) (Table 2). For maternal toenail arsenic, statistically significant negative association was observed among infants <20th percentile of birthweight distribution (Table 3). Conversely, among infants with birthweight above 2880g (45thpercentile), the association between arsenic and birthweight was in the positive direction, but not statistically significant. Traditional OLS based regression is unable to capture this underlying heterogeneity between arsenic-birthweight relations (Supplemental Table 1). Our results also indicate that via direct pathways, arsenic exposure will shift the birthweight distribution away from the mean. This change in birthweight distribution independent of gestational age can be expressed by a change in birthweight-for-gestational age Z-score, which is a transformation of birthweight that is independent of gestational age, in response to arsenic exposure (Supplemental Figure 1B, 1D). This suggests that prenatal arsenic exposure may be associated with decreased birthweight via intrauterine growth restriction among the smallest infants.

The direct and indirect effects were summed to obtain total effects of arsenic exposure on birthweight. Our results showed marked heterogeneity in the total effects of arsenic on birthweight across birthweight quantiles (figure 2E, 3E), which OLS regression did not capture (Supplemental Table 1). Overall, water arsenic exposure was negatively associated with birthweight among infants <45thpercentile of birthweight distribution, while the association was strongest among the smaller infants (Table 2). For instance, each unit increase in natural log water arsenic exposure was associated with 57g (95%CI: −83.6, −28.8) decrease in birthweight among infants with birthweight <2100g (5thpercentile) compared to 17.3g (95% CI: −35.0, −1.4) decrease in birthweight among infants with birthweight around 2820g (45thpercentile). Similar associations were observed for toenails arsenic (Table 3). Our results suggested that via both pathways, arsenic exposure was associated with a shift of the left slope of the birthweight

distribution curve away from the mean with a heavier lower tail (Supplemental Figure 1A, 1C). This change in the birthweight distribution curve was accompanied by a higher percentage of small preterm infants (≈9.2%) in our cohort compared to only 2-5% in a typical population.(Wilcox 2001)

Discussion

In our prospective cohort study, prenatal arsenic exposure was associated with decreased birthweight, and that the associations varied across birthweight percentiles. The effect of arsenic on birthweight was operated via pathways mediated through gestational age as well as pathways independent of gestational age, suggesting possible role of shortening of gestational age and intrauterine growth restriction in explaining arsenic-birthweight relations. The effect of arsenic on birthweight via shortening of gestational age negatively affected birthweight among all infants irrespective of birth sizes; although, the magnitude of the association was larger among infants at lower percentiles of birthweight distribution. On the other hand, the effect of arsenic on birthweight through pathways independent of gestational age negatively affected birthweight only among the smaller infants. Therefore, infants at the lower percentiles of birthweight distribution not only experienced decreased birthweight because of shortened gestation and intrauterine growth restriction due to arsenic exposure, but the magnitude of the associations were strongest among this group, suggesting heightened susceptibility to arsenic exposure among smaller infants, who already might be at higher perinatal risk. The overall negative effect of arsenic on birthweight observed among the smaller infants attenuated in magnitude and statistical significance as birthweight increased, until we observed non-significant positive associations among the heavier infants, suggesting that the susceptibility of arsenic exposure may vary by

infant's growth conditions. These associations were missed when traditional regression method was applied.

Our findings are consistent with previous epidemiological studies that also report that prenatal arsenic exposure is negatively associated with birthweight(Rahman et al. 2009, Kile et al. 2016, Huyck et al. 2007, Hopenhayn et al. 2003, Xu et al. 2011) and positively associated with preterm delivery,(Ahmad et al. 2001, Chakraborti et al. 2003, Yang et al. 2003) and fetal growth restriction.(Kippler et al. 2012, Thomas et al. 2015, Llanos and Ronco 2009) Our quantile regression results were also fairly consistent in magnitude and direction with mean regression analyses in the same cohort,(Kile et al. 2016) but captured additional shift in the birthweight distribution due to arsenic exposure. For instance, Kile et. al. previously estimated in this cohort that a unit increase in natural log drinking water and maternal toenail arsenic decreased birthweight by 17.4 g (-22.8, -12.0) and 13.6 g (-22.1, -5.1), respectively via decreasing gestational age,(Kile et al. 2016) which were comparable to the estimates we obtained at 50^{th} percentile of birthweight distribution (median regression) with increased sample size for water (n=1,140 vs 1,181) and toenail (n= 624 vs 1,104) arsenic exposure. Additionally, our analyses revealed a significant negative association between arsenic and birthweight among the smaller infants through pathways independent of gestational age that was not captured by OLS regression technique used by Kile et al.(Kile et al. 2016) This difference is likely due to linear modeling assumptions made by Kile et al.(Kile et al. 2016) The negative association between arsenic and birthweight independent of gestational age suggested that arsenic exposure might be associated with decreased birthweight among the smaller infants (i.e. $<20^{th}$–25^{th} percentiles) via growth restriction. This proportion of IUGR infants obtained using quantile causal mediation analysis technique was comparable to proportion of small-for-gestational age infants (≈19%),

which is a commonly accepted proxy measure for IUGR, in this cohort based on the global reference population.(Villar et al. 2014) Hence, our method could be useful to identify IUGR infants in a population without the use of any reference population.

While the heterogeneity in the associations between arsenic exposure and birthweight were identified in previous studies based on smoking,(Bloom et al. 2016) infant genders,(Xu et al. 2011, Gilbert-Diamond et al. 2016) and maternal prepregnancy BMI status,(Gilbert-Diamond et al. 2016) we observed disparities based on birthweight levels. The use of traditional OLS regression that estimates the change in mean birthweight summarizes effect estimates that differ across the range of the birthweight percentiles, including those with opposing signs and might underestimate the true public health burden of low birthweight associated with arsenic exposure. Our results indicated that prenatal arsenic exposure may be associated with a shift of the birthweight distribution curve away from the mean with a heavier lower tail, resulting in higher proportion of small preterm infants (≈9.2%) in our cohort. The proportion of small preterm infants in a population, which typically ranges between 2-5%, is an indicator of perinatal risk in that population.(Umbach and Wilcox 1996, Wilcox 2001) Our findings emphasized the importance of arsenic mitigation to improve perinatal outcomes in Bangladesh.

Known biologic effects of inorganic arsenic exposure support the biological plausibility of our findings. Arsenic can generate reactive oxygen species and deplete antioxidant enzymes (e.g. glutathione) leading to oxidative stress.(Jomova et al. 2011) Oxidative damage in early pregnancy can disrupt placental development, function and remodeling,(Jauniaux and Burton 2016) which in turn can hamper oxygen and nutrient supply to the growing fetus and production and metabolism of fetal growth regulating hormones leading to preterm delivery and IUGR.(Murphy et al. 2006, Parman, Wiley, and Wells 1999) Another plausible explanation is

epigenetic alterations. Prenatal arsenic exposure has been found associated with deregulation of microRNA expression profiles in umbilical cord blood(Rager et al. 2013) and DNA methylation status in maternal and umbilical cord blood.(Kile et al. 2012) MicroRNAs have important role in normal placental development; and alteration of microRNA expression profiles have been associated with abnormal placentation, preeclampsia, eclampsia, and SGA births.(Pineles et al. 2007, Zhu et al. 2009)

Our study has some limitations. The observed positive associations between arsenic exposure and birthweight among the heavier infants in our cohort could partly because of inadequate adjustment for maternal perinatal nutritional status. Individual's micronutrient status (e.g. folate, antioxidants) plays an important role in arsenic detoxification, where adequate nutrition may ameliorate individual's vulnerability to arsenic toxicity,(Milton et al. 2004) resulting in heavier infants. Therefore, inadequate adjustment for perinatal nutritional status will lead to an underestimation of the negative associations between arsenic and birthweight, which will be larger among heavier infants. For instance, we observed that the positive associations among the heavier infants were stronger in absence of maternal hemoglobin level in our models, which could be considered as a surrogate for perinatal nutritional status (results not shown). Future studies will be useful to explain potential interaction between prenatal arsenic exposure and maternal perinatal nutritional status on birthweight.

Hence, it is possible that there is error in our estimates due to unmeasured confounders. We selected a priori list of covariates that were previously found to be associated with our outcome (e.g. birthweight) and mediators (e.g. gestational age, and maternal hemoglobin) of interest. All women were provided with free prenatal multivitamins and the same level of health care during pregnancy by our community health clinics, which were among the few of healthcare

providers in that area. The compliance of regular multivitamins intake was reported to be 99%. We did not collect detail pregnancy history to adequately control for other factors that could confound our results, such as pregnancy spacing, or history of previous births. Furthermore, we were unable to test the robustness of our estimates in presence of unmeasured confounders because no such method had yet been developed for quantile causal mediation analyses technique. However, we tested the robustness of the indirect effect of arsenic exposure on birthweight that one might expect using traditional OLS regression considering gestational age as the lone mediator using sensitivity analyses technique.(Imai, Keele, and Tingley 2010) Our results indicated that if an unmeasured confounder effected both gestational age and birthweight in the same direction (e.g., perinatal care, birth defects, birth complications such as preeclampsia or eclampsia), then that particular factor would have to explain ≈12% of the original variance in gestational age and ≈8% of the original variance in birthweight for the causal mediation effect of arsenic on birthweight through gestational age to become zero (results not shown). On the other hand, if an unmeasured confounder affected gestational age and birthweight in the opposite direction (i.e. gestational diabetes), then failing to adjust for that factor would lead to an underestimation of the indirect effect via gestational age. These results indicate that our estimates for the causal mediation effect via gestation age are likely to be fairly robust to unmeasured confounding assumptions.

Strengths of our study include its prospective design, where we collected birth outcome data from a fairly large number of pregnant mothers. Arsenic exposure was measured in drinking water that pregnant mothers identified as their primary water source early in pregnancy and in maternal toenails collected within one month postpartum, reducing the risk of exposure misclassification. Maternal toenail arsenic represented individual's cumulative exposure over the

past 9−12 months,(Chen, Amarasiriwardena, and Christiani 1999) corresponding to entire pregnancy. Drinking water arsenic, which showed little temporal variability in previous studies,(Slotnick, Meliker, and Nriagu 2006) also served as an adequate marker for individual's long-term exposure, particularly when measured in main water sources.(Sohel et al. 2010) Therefore, our proposed temporal association for causal mediation framework between the exposure, mediator and outcome is properly framed. Gestational age was measured between 4−16 weeks of pregnancy using ultrasound technique, which is considered gold standard for pregnancy dating if used in early gestation.(Butt et al. 2014) Our quantile causal mediation analysis technique helped us to address methodological challenges involved in the investigation of potential heterogeneous associations between prenatal arsenic exposure and birthweight in relation to shortened gestation and intrauterine growth restriction. This novel modeling approach also allowed for potential exposure-mediator interaction,(Valeri and Vanderweele 2013) and correctly handled multiple mediators(VanderWeele and Vansteelandt 2014), which were advantages not shared by most of other mediation analysis approaches. Hence, our results are based on fairly solid ground.

Conclusions

Our results showed that prenatal arsenic exposure was associated with decreased birthweight and that the magnitude of the association varied across birthweight percentiles. The smaller infants, who were already at high risk of perinatal mortality and morbidity, were more susceptible to the negative effect of arsenic on birthweight because of both shortening of gestational age and intrauterine growth restriction. Minimizing maternal arsenic exposure during pregnancy may significantly improve perinatal outcomes in Bangladesh.

Bibliography

Ahamed, S., M. K. Sengupta, S. C. Mukherjee, S. Pati, A. Mukherjeel, M. M. Rahman, M. A. Hossain, B. Das, B. Nayakl, A. Pal, A. Zafar, S. Kabir, S. A. Banu, S. Morshed, T. Islam, M. M. Rahman, Q. Quamruzzaman, and D. Chakraborti. 2006. "An eight-year study report on arsenic contamination in groundwater and health effects in Eruani village, Bangladesh and an approach for its mitigation." *J Health Popul Nutr* 24 (2):129-41.

Ahmad, S. A., M. H. Sayed, S. Barua, M. H. Khan, M. H. Faruquee, A. Jalil, S. A. Hadi, and H. K. Talukder. 2001. "Arsenic in drinking water and pregnancy outcomes." *Environ Health Perspect* 109 (6):629-31.

Barker, D. J., J. G. Eriksson, T. Forsen, and C. Osmond. 2002. "Fetal origins of adult disease: strength of effects and biological basis." *Int J Epidemiol* 31 (6):1235-9.

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

Butt, K., K. Lim, Obstetricians Society of, and Canada Gynaecologists of. 2014. "Determination of gestational age by ultrasound." *J Obstet Gynaecol Can* 36 (2):171-83.

Chakraborti, D., S. C. Mukherjee, S. Pati, M. K. Sengupta, M. M. Rahman, U. K. Chowdhury, D. Lodh, C. R. Chanda, A. K. Chakraborti, and G. K. Basu. 2003. "Arsenic groundwater contamination in Middle Ganga Plain, Bihar, India: a future danger?" *Environ Health Perspect* 111 (9):1194-201.

Chakraborti, D., M. K. Sengupta, M. M. Rahman, S. Ahamed, U. K. Chowdhury, M. A. Hossain, S. C. Mukherjee, S. Pati, K. C. Saha, R. N. Dutta, and Q. Quamruzzaman. 2004. "Groundwater arsenic contamination and its health effects in the Ganga-Meghna-Brahmaputra plain." *J Environ Monit* 6 (6):74N-83N.

Chen, K. L., C. J. Amarasiriwardena, and D. C. Christiani. 1999. "Determination of total arsenic concentrations in nails by inductively coupled plasma mass spectrometry." *Biol Trace Elem Res* 67 (2):109-25. doi: 10.1007/BF02784067.

Claus Henn, B., A. S. Ettinger, M. R. Hopkins, R. Jim, C. Amarasiriwardena, D. C. Christiani, B. A. Coull, D. C. Bellinger, and R. O. Wright. 2016. "Prenatal Arsenic Exposure and Birth Outcomes among a Population Residing near a Mining-Related Superfund Site." *Environ Health Perspect* 124 (8):1308-15. doi: 10.1289/ehp.1510070.

Concha, G., B. Nermell, and M. V. Vahter. 1998. "Metabolism of inorganic arsenic in children with chronic high arsenic exposure in northern Argentina." *Environ Health Perspect* 106 (6):355-9.

Fei, D. L., D. C. Koestler, Z. Li, C. Giambelli, A. Sanchez-Mejias, J. A. Gosse, C. J. Marsit, M. R. Karagas, and D. J. Robbins. 2013. "Association between In Utero arsenic exposure, placental gene expression, and infant birth weight: a US birth cohort study." *Environ Health* 12:58. doi: 10.1186/1476-069X-12-58.

Flanagan, S. V., R. B. Johnston, and Y. Zheng. 2012. "Arsenic in tube well water in Bangladesh: health and economic impacts and implications for arsenic mitigation." *Bull World Health Organ* 90 (11):839-46. doi: 10.2471/BLT.11.101253.

Gelmann, E. R., E. Gurzau, A. Gurzau, W. Goessler, J. Kunrath, C. W. Yeckel, and K. M. McCarty. 2013. "A pilot study: the importance of inter-individual differences in inorganic arsenic metabolism for birth weight outcome." *Environ Toxicol Pharmacol* 36 (3):1266-75. doi: 10.1016/j.etap.2013.10.006.

Gilbert-Diamond, D., J. A. Emond, E. R. Baker, S. A. Korrick, and M. R. Karagas. 2016. "Relation between in Utero Arsenic Exposure and Birth Outcomes in a Cohort of Mothers and Their Newborns from New Hampshire." *Environ Health Perspect* 124 (8):1299-307. doi: 10.1289/ehp.1510065.

Guan, H., F. Piao, X. Zhang, X. Li, Q. Li, L. Xu, F. Kitamura, and K. Yokoyama. 2012. "Prenatal exposure to arsenic and its effects on fetal development in the general population of Dalian." *Biol Trace Elem Res* 149 (1):10-5. doi: 10.1007/s12011-012-9396-7.

He, W., R. J. Greenwell, D. M. Brooks, L. Calderon-Garciduenas, H. D. Beall, and J. D. Coffin. 2007. "Arsenic exposure in pregnant mice disrupts placental vasculogenesis and causes spontaneous abortion." *Toxicol Sci* 99 (1):244-53. doi: 10.1093/toxsci/kfm162.

Hellman, L. M., M. Kobayashi, L. Fillisti, M. Lavenhar, and E. Cromb. 1969. "Growth and development of the human fetus prior to the twentieth week of gestation." *Am J Obstet Gynecol* 103 (6):789-800.

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

Hosain, G. M., N. Chatterjee, A. Begum, and S. C. Saha. 2006. "Factors associated with low birthweight in rural Bangladesh." *J Trop Pediatr* 52 (2):87-91. doi: 10.1093/tropej/fmi066.

Hu, G., S. Bi, G. Xu, Y. Zhang, X. Mei, and A. Li. 2015. "Distribution and assessment of heavy metals off the Changjiang River mouth and adjacent area during the past century and the relationship of the heavy metals with anthropogenic activity." *Mar Pollut Bull* 96 (1-2):434-40. doi: 10.1016/j.marpolbul.2015.05.009.

Huyck, K. L., M. L. Kile, G. Mahiuddin, Q. Quamruzzaman, M. Rahman, C. V. Breton, C. B. Dobson, J. Frelich, E. Hoffman, J. Yousuf, S. Afroz, S. Islam, and D. C. Christiani. 2007. "Maternal arsenic exposure associated with low birth weight in Bangladesh." *J Occup Environ Med* 49 (10):1097-104. doi: 10.1097/JOM.0b013e3181566ba0.

Imai, K., L. Keele, and D. Tingley. 2010. "A general approach to causal mediation analysis." *Psychol Methods* 15 (4):309-34. doi: 10.1037/a0020761.

Jauniaux, E., and G. J. Burton. 2016. "[The role of oxidative stress in placental-related diseases of pregnancy]." *J Gynecol Obstet Biol Reprod (Paris)* 45 (8):775-785. doi: 10.1016/j.jgyn.2016.02.012.

Jomova, K., Z. Jenisova, M. Feszterova, S. Baros, J. Liska, D. Hudecova, C. J. Rhodes, and M. Valko. 2011. "Arsenic: toxicity, oxidative stress and human disease." *J Appl Toxicol* 31 (2):95-107. doi: 10.1002/jat.1649.

Kile, M. L., A. Baccarelli, E. Hoffman, L. Tarantini, Q. Quamruzzaman, M. Rahman, G. Mahiuddin, G. Mostofa, Y. M. Hsueh, R. O. Wright, and D. C. Christiani. 2012. "Prenatal arsenic exposure and DNA methylation in maternal and umbilical cord blood leukocytes." *Environ Health Perspect* 120 (7):1061-6. doi: 10.1289/ehp.1104173.

Kile, M. L., A. Cardenas, E. Rodrigues, M. Mazumdar, C. Dobson, M. Golam, Q. Quamruzzaman, M. Rahman, and D. C. Christiani. 2016. "Estimating Effects of Arsenic Exposure During Pregnancy on Perinatal Outcomes in a Bangladeshi Cohort." *Epidemiology* 27 (2):173-81. doi: 10.1097/EDE.0000000000000416.

Kile, M. L., E. G. Rodrigues, M. Mazumdar, C. B. Dobson, N. Diao, M. Golam, Q. Quamruzzaman, M. Rahman, and D. C. Christiani. 2014. "A prospective cohort study of the association between drinking water arsenic exposure and self-reported maternal health symptoms during pregnancy in Bangladesh." *Environ Health* 13 (1):29. doi: 10.1186/1476-069X-13-29.

Kippler, M., Y. Wagatsuma, A. Rahman, B. Nermell, L. A. Persson, R. Raqib, and M. Vahter. 2012. "Environmental exposure to arsenic and cadmium during pregnancy and fetal size: a longitudinal study in rural Bangladesh." *Reprod Toxicol* 34 (4):504-11. doi: 10.1016/j.reprotox.2012.08.002.

Kline, Jennie; Susser, Mervyn; Stein, Zena. 1989. *Conception to birth : epidemiology of prenatal development*. Vol. 14, *Monographs in epidemiology and biostatistics*. New York: Oxford University Press.

Kramer, M. S. 1987. "Determinants of low birth weight: methodological assessment and meta-analysis." *Bull World Health Organ* 65 (5):663-737.

Kwok, R. K., R. B. Kaufmann, and M. Jakariya. 2006. "Arsenic in drinking-water and reproductive health outcomes: a study of participants in the Bangladesh Integrated Nutrition Programme." *J Health Popul Nutr* 24 (2):190-205.

Laine, J. E., K. A. Bailey, M. Rubio-Andrade, A. F. Olshan, L. Smeester, Z. Drobna, A. H. Herring, M. Styblo, G. G. Garcia-Vargas, and R. C. Fry. 2015. "Maternal arsenic exposure, arsenic methylation efficiency, and birth outcomes in the Biomarkers of Exposure to ARsenic (BEAR) pregnancy cohort in Mexico." *Environ Health Perspect* 123 (2):186-92. doi: 10.1289/ehp.1307476.

Lee, A. C., J. Katz, H. Blencowe, S. Cousens, N. Kozuki, J. P. Vogel, L. Adair, A. H. Baqui, Z. A. Bhutta, L. E. Caulfield, P. Christian, S. E. Clarke, M. Ezzati, W. Fawzi, R. Gonzalez, L. Huybregts, S. Kariuki, P. Kolsteren, J. Lusingu, T. Marchant, M. Merialdi, A. Mongkolchati, L. C. Mullany, J. Ndirangu, M. L. Newell, J. K. Nien, D. Osrin, D. Roberfroid, H. E. Rosen, A. Sania, M. F. Silveira, J. Tielsch, A. Vaidya, B. A. Willey, J. E. Lawn, R. E. Black, and Cherg Sga- Preterm Birth Working Group. 2013. "National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010." *Lancet Glob Health* 1 (1):e26-36. doi: 10.1016/S2214-109X(13)70006-8.

Llanos, M. N., and A. M. Ronco. 2009. "Fetal growth restriction is related to placental levels of cadmium, lead and arsenic but not with antioxidant activities." *Reprod Toxicol* 27 (1):88-92. doi: 10.1016/j.reprotox.2008.11.057.

Loewenberg, S. 2016. "In Bangladesh, arsenic poisoning is a neglected issue." *Lancet* 388 (10058):2336-2337. doi: 10.1016/S0140-6736(16)32173-0.

McDermott, S., W. Bao, C. M. Aelion, B. Cai, and A. B. Lawson. 2014. "Does the metal content in soil around a pregnant woman's home increase the risk of low birth weight for her infant?" *Environ Geochem Health* 36 (6):1191-7. doi: 10.1007/s10653-014-9617-4.

Milton, A. H., Z. Hasan, S. M. Shahidullah, S. Sharmin, M. D. Jakariya, M. Rahman, K. Dear, and W. Smith. 2004. "Association between nutritional status and arsenicosis due to chronic arsenic exposure in Bangladesh." *Int J Environ Health Res* 14 (2):99-108. doi: 10.1080/0960312042000209516.

Milton, A. H., W. Smith, B. Rahman, Z. Hasan, U. Kulsum, K. Dear, M. Rakibuddin, and A. Ali. 2005. "Chronic arsenic exposure and adverse pregnancy outcomes in bangladesh." *Epidemiology* 16 (1):82-6.

Mukherjee, S. C., K. C. Saha, S. Pati, R. N. Dutta, M. M. Rahman, M. K. Sengupta, S. Ahamed, D. Lodh, B. Das, M. A. Hossain, B. Nayak, A. Mukherjee, D. Chakraborti, S. K. Dulta, S. K. Palit, I. Kaies, A. K. Barua, and K. A. Asad. 2005. "Murshidabad--one of the nine groundwater arsenic-affected districts of West Bengal, India. Part II: dermatological, neurological, and obstetric findings." *Clin Toxicol (Phila)* 43 (7):835-48.

Murphy, V. E., R. Smith, W. B. Giles, and V. L. Clifton. 2006. "Endocrine regulation of human fetal growth: the role of the mother, placenta, and fetus." *Endocr Rev* 27 (2):141-69. doi: 10.1210/er.2005-0011.

Myers, S. L., D. T. Lobdell, Z. Liu, Y. Xia, H. Ren, Y. Li, R. K. Kwok, J. L. Mumford, and P. Mendola. 2010. "Maternal drinking water arsenic exposure and perinatal outcomes in inner Mongolia, China." *J Epidemiol Community Health* 64 (4):325-9. doi: 10.1136/jech.2008.084392.

Parman, T., M. J. Wiley, and P. G. Wells. 1999. "Free radical-mediated oxidative DNA damage in the mechanism of thalidomide teratogenicity." *Nat Med* 5 (5):582-5. doi: 10.1038/8466.

Pineles, B. L., R. Romero, D. Montenegro, A. L. Tarca, Y. M. Han, Y. M. Kim, S. Draghici, J. Espinoza, J. P. Kusanovic, P. Mittal, S. S. Hassan, and C. J. Kim. 2007. "Distinct subsets of microRNAs are expressed differentially in the human placentas of patients with preeclampsia." *Am J Obstet Gynecol* 196 (3):261 e1-6. doi: 10.1016/j.ajog.2007.01.008.

Rager, J. E., K. A. Bailey, L. Smeester, S. K. Miller, J. S. Parker, J. E. Laine, Z. Drobna, J. Currier, C. Douillet, A. F. Olshan, M. Rubio-Andrade, M. Styblo, G. Garcia-Vargas, and R. C. Fry. 2013. "Prenatal arsenic exposure and the epigenome: Altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood." *Environ Mol Mutagen*. doi: 10.1002/em.21842.

Rahman, A., L. A. Persson, B. Nermell, S. El Arifeen, E. C. Ekstrom, A. H. Smith, and M. Vahter. 2010. "Arsenic exposure and risk of spontaneous abortion, stillbirth, and infant mortality." *Epidemiology* 21 (6):797-804. doi: 10.1097/EDE.0b013e3181f56a0d.

Rahman, A., M. Vahter, A. H. Smith, B. Nermell, M. Yunus, S. El Arifeen, L. A. Persson, and E. C. Ekstrom. 2009. "Arsenic exposure during pregnancy and size at birth: a prospective cohort study in Bangladesh." *Am J Epidemiol* 169 (3):304-12. doi: 10.1093/aje/kwn332.

Robinson, H. P., and J. E. Fleming. 1975. "A critical evaluation of sonar "crown-rump length" measurements." *Br J Obstet Gynaecol* 82 (9):702-10.

Shirai, S., Y. Suzuki, J. Yoshinaga, and Y. Mizumoto. 2010. "Maternal exposure to low-level heavy metals during pregnancy and birth size." *J Environ Sci Health A Tox Hazard Subst Environ Eng* 45 (11):1468-74. doi: 10.1080/10934529.2010.500942.

Slotnick, M. J., J. R. Meliker, and J. O. Nriagu. 2006. "Effects of time and point-of-use devices on arsenic levels in Southeastern Michigan drinking water, USA." *Sci Total Environ* 369 (1-3):42-50. doi: 10.1016/j.scitotenv.2006.04.021.

Sohel, N., P. S. Kanaroglou, L. A. Persson, M. Z. Haq, M. Rahman, and M. Vahter. 2010. "Spatial modelling of individual arsenic exposure via well water: evaluation of arsenic in urine, main water source and influence of neighbourhood water sources in rural Bangladesh." *J Environ Monit* 12 (6):1341-8. doi: 10.1039/c001708f.

Thomas, S., T. E. Arbuckle, M. Fisher, W. D. Fraser, A. Ettinger, and W. King. 2015. "Metals exposure and risk of small-for-gestational age birth in a Canadian birth cohort: The MIREC study." *Environ Res* 140:430-9. doi: 10.1016/j.envres.2015.04.018.

Umbach, D. M., and A. J. Wilcox. 1996. "A technique for measuring epidemiologically useful features of birthweight distributions." *Stat Med* 15 (13):1333-48. doi: 10.1002/(SICI)1097-0258(19960715)15:13<1333::AID-SIM271>3.0.CO;2-R.

Valeri, L., and T. J. Vanderweele. 2013. "Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros." *Psychol Methods* 18 (2):137-50. doi: 10.1037/a0031034.

Vall, O., M. Gomez-Culebras, O. Garcia-Algar, X. Joya, D. Velez, E. Rodriguez-Carrasco, and C. Puig. 2012. "Assessment of prenatal exposure to arsenic in Tenerife Island." *PLoS One* 7 (11):e50463. doi: 10.1371/journal.pone.0050463.

VanderWeele, T. J., and S. Vansteelandt. 2014. "Mediation Analysis with Multiple Mediators." *Epidemiol Method* 2 (1):95-115. doi: 10.1515/em-2012-0010.

Villar, J., L. Cheikh Ismail, C. G. Victora, E. O. Ohuma, E. Bertino, D. G. Altman, A. Lambert, A. T. Papageorghiou, M. Carvalho, Y. A. Jaffer, M. G. Gravett, M. Purwar, I. O. Frederick, A. J. Noble, R. Pang, F. C. Barros, C. Chumlea, Z. A. Bhutta, S. H. Kennedy, Fetal International, and Century Newborn Growth Consortium for the 21st. 2014. "International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project." *Lancet* 384 (9946):857-68. doi: 10.1016/S0140-6736(14)60932-6.

WHO and UNICEF, World Health Organization. 2004. "Low birthweight: country, regional and global estimates." accessed July 15, 2016. http://www.unicef.org/publications/files/low_birthweight_from_EY.pdf.

Wilcox, A. J. 2001. "On the importance--and the unimportance--of birthweight." *Int J Epidemiol* 30 (6):1233-41.

Xu, L., K. Yokoyama, Y. Tian, F. Y. Piao, F. Kitamura, H. Kida, and P. Wang. 2011. "Decrease in birth weight and gestational age by arsenic among the newborn in Shanghai, China." *Nihon Koshu Eisei Zasshi* 58 (2):89-95.

Yang, C. Y., C. C. Chang, S. S. Tsai, H. Y. Chuang, C. K. Ho, and T. N. Wu. 2003. "Arsenic in drinking water and adverse pregnancy outcome in an arseniasis-endemic area in northeastern Taiwan." *Environ Res* 91 (1):29-34.

Zhu, X. M., T. Han, I. L. Sargent, G. W. Yin, and Y. Q. Yao. 2009. "Differential expression profile of microRNAs in human placentas from preeclamptic pregnancies vs normal pregnancies." *Am J Obstet Gynecol* 200 (6):661 e1-7. doi: 10.1016/j.ajog.2008.12.045.

3. Chapter 3: Epigenetic regulation of birthweight in relation to environmental arsenic exposure: role of placenta-derived microRNAs

Authors: Mohammad Rahman[1], Liming Liang[6], Linda Valeri[2], Zhaozhong Zhu[1], Li Su[1], Shangzhi Gao[1], Golam Mostofa[4], Qazi Qamruzzaman[4], Andrea Baccarelli[5], Russ Hauser[1,7], David C. Christiani[1,7]

Author Affiliations:

[1]Harvard T.H. Chan School of Public Health, Department of Environmental Health, Boston, Massachusetts, USA

[2]McLean Hospital, Belmont, Massachusetts, USA and Harvard Medical School, Boston, Massachusetts, USA

[3] Oregon State University, College of Public Health and Human Sciences, Corvallis, Oregon, USA

[4]Dhaka Community Hospital Trust, Dhaka, Bangladesh

[5]Columbia University, Mailman School of Public Health, Department of Environmental Health, New York, NY, USA

[6]Harvard T.H. Chan School of Public Health, Department of Biostatistics, Boston, Massachusetts, USA

[7]Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, Massachusetts, USA

Abstract

Background: Dysregulation of the placental epigenome due to perturbations in the *in-utero* environment may influence fetal growth and development that have long-term health consequences. Since the placenta is the principal regulator of the *in-utero* environment, it expresses numerous microRNAs (miRNAs) that may be sensitive to environmental toxicants and therefore, associated with birth outcomes.

Methods: Placental samples from 77 cases and 77 controls, selected respectively from the lower and upper tails of birthweight distribution of 1,141 liveborn singletons from a longitudinal birth cohort in Bangladesh, were screened for 754 miRNAs using TaqMan qRT-PCR OpenArray. Expression of 49 miRNAs associated with the case-control status, or gestational age, or both were confirmed among 364 randomly selected participants using TaqMan pRT-PCR. Arsenic was measured in cord blood and gestational age was determined by ultrasound. Causal mediation analysis was used to estimate direct, indirect and total effects of 1) miRNA expressions on birthweight considering gestational age a mediator, and 2) *in-utero* arsenic exposure on birthweight considering miRNA expressions and gestational age as sequential mediators.

Results: Placental expression of miR-1290, and miR-195 showed negative associations with birthweight, while miR-328, and miR-324-5p showed positive associations (FDR q-value<0.05) via changing gestational age. The effect of miR-1290, miR-195, miR-328, and miR-489 on birthweight varied by gestational age, where the associations were stronger in magnitude and statistical significance in shorter pregnancies. miR-1290 and miR-195 also modulated the negative associations between arsenic and birthweight, where stronger associations were observed for higher expressions of these miRNAs, when the gestational age is also lower. *In*

silico functional analysis of predicted targets revealed several biological processes, including insulin signaling, cell proliferation, inflammation, and apoptosis associated with these miRNAs.

Conclusions: Placenta-derived miRNAs regulate birthweight by not only determining the length of gestation but also modulating the susceptibility of adverse extrauterine factors on fetal growth. Future studies to validate our study findings in an external cohort will be helpful.

Introduction

There is a growing interest in determining how prenatal exposure to environmental toxicants alter the programming of fetal and childhood growth and development that have long-term consequences on later-life health and disease risk. Low birthweight is one of the earliest phenotypic manifestations of altered fetal programming resulting from perturbations in the intrauterine environment due to factors related to the extrauterine environment, including environmental arsenic exposure. Low birthweight has been linked with mortality and morbidity in early infancy as well as increased risk of certain diseases later in life, particularly adult-onset diabetes mellitus, metabolic abnormalities, obesity, coronary heart disease, hypertension, and stroke.(Barker 2004) Identifying *in-utero* molecular markers that are associated with low birthweight and environmental arsenic exposure may provide clues to understand fetal programming and its role in altering life-long health trajectories.

The placenta is the principal regulator of the intrauterine environment, maintaining fetal growth and development by facilitating fetal supply of nutrients and oxygen, secreting hormones, and, detoxifying potentially harmful maternal factors and environmental toxins.(Gude et al. 2004) Exogenous environmental toxicants, such as inorganic arsenic, have been shown to readily cross and accumulate in the placenta and disrupt the regulatory function of the placenta(Concha, Nermell, and Vahter 1998). But, the mechanisms by which intrauterine environment dysregulate

the normal homoeostatic response of the placenta are not well understood. One possible explanation is the re-programming of placental epigenetic patterns, including altered expression of microRNAs (miRNAs) by the placenta.(Marsit 2015)

miRNAs are 18-25 nucleotide long small non-coding RNA molecules involved in post-transcriptional gene regulation.(Lee, Feinbaum, and Ambros 1993) MicroRNA binds to the 3'-untranslated region of the targeted messenger RNA (mRNA) and effectively silence gene expression by either repressing mRNA translation or by degrading the mRNA transcript.(Lee, Feinbaum, and Ambros 1993) Thus, miRNAs regulate nearly 30 to 60% of all human protein-coding genes and virtually every cellular process including apoptosis, cell proliferation, and cellular differentiation.(Calin and Croce 2006, Rager et al. 2013, Hagen and Lai 2008) miRNAs also exhibit tissue-specific expression and function, and numerous miRNAs are specifically expressed in the placenta.(Bentwich et al. 2005) Altered expression of miRNAs in the placenta and cord blood has also been associated with prenatal exposure to environmental toxicants such as arsenic, biphenyl-A, and smoking, (Maccani et al. 2010, Rager et al. 2013) as well as adverse pregnancy outcomes, including abnormal placentation, preeclampsia or eclampsia, and small for gestational age birth.(Pineles et al. 2007, Zhu et al. 2009) Yet, few studies that looked at the associations between miRNA expressions in the placenta and birthweight frequently adopted a candidate miRNA expression analysis approach with limited study participants, and often in absence of an exposure of interest.

In this study, we aim to identify placenta-derived miRNA markers that are associated with birthweight using global miRNA expression analysis approach, and further characterize the role of placenta-derived miRNAs on birthweight in response to environmental arsenic exposure in a longitudinal birth cohort in Bangladesh. We hypothesized that epigenetic alterations

responsible for abnormal fetal growth and development will be captured by miRNA expression profiles in the placenta, and that prenatal exposure to environmental toxicants may interfere with the homeostatic epigenetic regulation of birthweight by the placenta.

Methods

Study population and subject selection

We applied a two-phase study with total 458 participants sampled from a longitudinal birth cohort in Bangladesh. During 2008-2011, 1,613 pregnant mothers were enrolled from Pabna and Sirajdikhan Upazilas of Bangladesh with reportable exposure to inorganic arsenic through contaminated drinking water in order to study arsenic induced reproductive health outcomes. The details of this study, including recruitment and enrollment criteria were previously reported(Kile et al. 2014). Briefly, women aged 18 years or older with an ultrasound confirmed singleton pregnancy of ≤16 weeks' gestation were recruited and followed throughout pregnancy. At the end of follow-up, 1,184 singleton livebirths were recorded after exclusions due to loss of contact before delivery (n=99), participation withdrawn (n=121), miscarriage (n=132), stillbirth (n=72), and twin births (n=5). Tissue samples from the placenta were available from 1,141 participants.

In the first phase of the study, we employed extreme phenotype sampling to select a discovery cohort of 154 participants: 77 cases with birth weight ranging from 800-2200 gm and 77 controls with birth weight ranging from 3300-4000 gm. The extreme phenotype sampling in the discovery cohort was designed for screening purpose in order to maximize the predictive power. Candidate microRNA selected from the discovery cohort were then replicated in a cohort of 364 participants randomly selected from the original cohort with available placental tissue samples. Thus, 52 participants included in the discovery cohort were also included in the

replication cohort, allowing us to assess between platform variability in miRNA expressions. More details of the study design and analysis approach appear in Supplemental Table 1.

Outcome and covariates

Birth weight was measured on a pediatric scale calibrated before each measurement and rounded to the nearest 10 grams. Gestational age was determined by ultrasonography by using crown-rump length at median 12 [range: 4–16] weeks of pregnancy. We used a structured questionnaire to collect socio-demographic, medical, and environmental information including secondhand smoke, number of past pregnancy, household income. Trained health care workers measured maternal height, weight and collected blood for hemoglobin concentration at the time of enrollment. Health care workers made monthly house visits to distribute prenatal vitamins, record height and weight, and measure blood pressure.

Exposure assessment

Exposure to inorganic arsenic was measured in cord blood collected in trace metal-free vacutainer tube containing K2EDTA (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) and stored at 4 °C until it was shipped to Harvard TH Chan School of Public Health (HSPH) for storage. Cord blood was analyzed at the Trace Metals Laboratory at the HSPH. Umbilical cord blood samples were first weighed (~1 g) and digested for 24 h in 2 mL of concentrated nitric acid. The samples were then treated with 1 mL of 30 % hydrogen peroxide per 1 g of blood and left overnight. Samples were subsequently diluted to 10 mL with deionized water. Blood lead concentrations were measured using a dynamic reaction cell-inductively coupled plasma mass spectrometer (DRC-ICP-MS, DRC II, Perkin Elmer). Continual calibration standards and a standard reference solution (NIST 1643e: Metals in Water) were used to monitor precision and accuracy of the analysis. The percent recovery rate for inorganic ranged from 81 to

108 %, and precision was measured as percent relative standard deviation (SD), with a result of less than 3 % for arsenic. The average limit of detection (LOD) was 0.03 μg/dL. For sample concentrations below the LOD (n=3), the instrumental values were used for analysis.

Placental tissue collection

Placental tissue was sampled from the fetal side between 4-120 minutes after delivery. Two specimens, one from the inner region proximal to the umbilical cord insertion point and one from the outer region closer to the edge, were collected from the placenta to control for within-placenta variability. Methods and rationale for the sampling scheme were described previously (Adibi et al. 2009). Careful dissection was used to maximize the amount of villous tissue in the sample and avoid membrane and decidual contamination. Samples were preserved in RNAlater (Ambion, Austin, TX) to stabilize the RNA and shipped at Harvard T.H. Chan School of Public Health, where they was stored at −80°C.

RNA isolation

Total RNA containing small RNA was extracted from 50 mg of thawed placenta. The tissue was homogenized in QIAzol Lysis Reagent followed by RNA isolation using QIAgen miRNeasy kit following manufacturer's guidelines (Qiagen, Valencia, CA, USA). Agilient 2100 Bioanalyzer (Agilient Technologies, Palo Alto, CA) was used to assess RNA quality. RNA was transcribed into cDNA in 2 multiplex reactions each containing 3 μl of the small RNA preparation and either TaqMan Megaplex RT primers, Human Pool A or Pool B and then preamplified using TaqMan Megaplex PreAmp Primers, Human Pool A and Pool B. An amount of 200 ng of total RNA containing the small RNA fraction was used for each sample for reverse transcription with Megaplex^{TM}

MicroRNA expression analysis

All miRNA analyses were performed at the Genetic Resources Core Facility, the Johns Hopkins School of Medicine, Institute of Genetic Medicine, Baltimore, Maryland, USA.

Discovery phase: MicroRNA expression analysis in the discovery cohort used TaqMan OpenArray Human MicroRNA Panel (Applied Biosystems, Foster City, CA) with 754 preconfigured human miRNAs following manufacturer's guidelines. Each plate contained 3 TaqMan miRNA assay endogenous controls and one TaqMan miRNA assay not related to human as negative controls to aid in data normalization. The cDNA was run through 40 cycles of 16^{o} C for 2 minutes, 42^{o} C for 1 minute, and 50^{o} C for 1 second. Reactions were then held at 85^{o}C for 5 minutes and cooled to 4^{o} C for storage.

Replication phase: 52 miRNAs, that included 3 internal control miRNA (miR-362, miR-532, miR-532-3p) were selected from the discovery phase and analyzed by using customized TaqMan OpenArray real-time polymerase chain system. Out of 364 replication samples, 56 samples were analyzed in duplicates, and 310 samples were analyzed in triplicates. We used technical duplicates in order to reduce confounding and bias from technical variations within and across the phases. The quality of qRT-PCR data between the discovery and replication phases appears in supplemental Table 3.2.

Statistical analysis

Discovery Phase: The population characteristics between the case and control groups were assessed by ANOVA, Wilcoxon rank sum, or Chi-square tests. Data were checked for outliers using box plots and samples with low RNA yield or absorbance were excluded. One sample with very few detectable miRNAs was excluded. We used the following quality control criteria in order to identify reliable miRNAs: high expression measured by cycle threshold (Ct) <30; good amplification quality measured by amplification score >1.1 and Cq confidence >0.8;

and detectable expression in more than 50% samples among the cases or controls. 306 miRNAs passed quality control criteria and were used for statistical analysis.

We used two different imputation strategies to analyze missing values effect. First, we imputed missing values for a specific miRNA for each samples among the cases and controls with the corresponding 95^{th} percentile for the same miRNA among cases and controls without missing values. This strategy makes the assumption that Ct values were not missing at random but missing due to no or very low expression of that miRNA. Second, we used multiple imputations by fitting a sequence of regression models and drawing values from the distribution of all miRNAs under the assumption that missing values were at random and were a result of failed measurement.(Raghunathan TE 2001) Test statistics of all miRNAs for both imputed strategies were compared with the test statistics obtained by the complete case analysis.

To test the sensitivity of results due to technical variations, analysis was conducted on raw cycle threshold values as well as normalized values. Four normalization methods: global normalization, quantile normalization, endogene normalization, and keep as raw were applied using R/Bioconductor to verify the sensitivity of results. For endogene normalization, geometric averaging of three most stable internal control miRNAs with less than 3% missing rates was identified and used for normalization following the method described by Vendesompele et al. using R/Bioconductor (D'Haene et al. 2012, Vandesompele et al. 2002) Statistical analysis was applied separately for each normalization methods.

Known or unknown biological or technical variations including batch effects may introduce variations in the measurement of transcript abundance in high-throughput experiments.(Akey et al. 2007) Hence, we constructed surrogate variables using information among hundreds of miRNAs on the OpenArray platform and used them in subsequent analysis to

adjust for unknown, unmodeled or latent sources of noise (Leek and Storey 2007). The association between miRNA expressions and the case-control status was assessed by means of multivariate linear regression models, where miRNA expressions (continuous) were used as the dependent variable and the case-control status (binary) as an independent variable in order to maximize the power. Based on our previous knowledge that gestational age may act as an intermediate or mediating variable between miRNA expressions and birthweight, we fit two sets of models- models for birthweight 1) adjusting for gestational age, and 2) not adjusting for gestational age. In addition, we fit a 3rd model to screen miRNAs associated with gestational age by means of a multivariate linear regression model adjusting for the case-control status. This approach allowed us to screen for miRNAs that might be associated with the case-control status either via gestational age (indirect pathway) or, via pathways independent of gestational age (direct pathway) or, via both pathways. Other covariates included in the models were study site, maternal age, and infant gender, which were selected a-priori based on our previous knowledge.

A separate model was fit for each miRNA and an observed *P*-value was generated. A permutation-based *P*-value was then calculated to assess significance. Briefly, we created two separate datasets- one for the miRNA expressions and the other for the phenotypic information. We then permuted 1000 phenotypic datasets by randomly shuffling the case control status, while keeping all other phenotypic information unchanged. Implementing the models described above, we then calculated 1000 p-values from the 1000 permuted datasets for each miRNA. The final P-value assigned to a miRNA was the fraction of permuted *P*-values equal to or more extreme than the observed *P*-value estimated in the parent dataset.

The selection of candidate miRNAs from the discovery cohort was based on the following criteria: permutation *P*-value less than 0.01 and 1) miRNA reached to significance

level in at least two out of four normalization methods; or 2) miRNA reached to significance level in any normalization method but have biological significance.

Replication phase: Data were checked for outliers, and samples with low expression due to technical problems were excluded from further analysis. As the top-hit miRNAs selected in the "Discovery Study" were analyzed either in duplicates (n=56) or triplicates (n=310) in the replication cohort, we measured the mean Ct value of each miRNA for each sample for statistical analysis. The same data quality control criteria as the "Discovery Study" were applied in the replication study. Data normalization was carried out using the same three internal control miRNAs (e.g. miR-362, miR-532, miR-532-3p) as in the "Discovery Study". Surrogate variables were also constructed to control for known, unknown technical and biological variations in miRNA expressions.

The associations between miRNA expression and birthweight were analyzed using causal mediation analysis approach considering gestational age a mediator (Valeri and Vanderweele 2013) The directed acyclic graph (DAG) in Figure 3.1 explains our conceptual model for causal mediation analysis. We hypothesized that gestational age will lie within the causal pathway between miRNA expressions and birthweight. Basically, we fit two sets of models- model for gestational age (the mediator model) and model for birthweight (the outcome model). Multivariate linear regressions were used to model for both gestational age and birthweight. Covariates included in the models were study site, maternal age, education, enrollment BMI, number of past pregnancy, infant

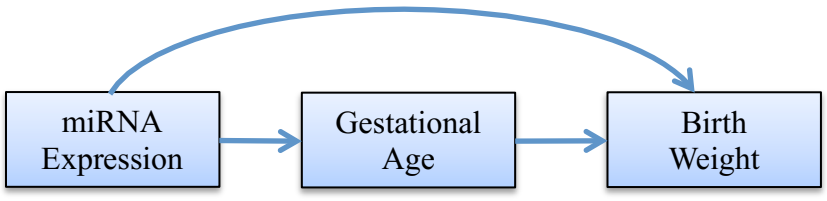

Figure 3.1. Directed acyclic graph showing conceptual model for causal mediation analysis of miRNA expressions and birthweight considering gestational age a mediator

gender, and surrogate variables. The model for birthweight was additionally adjusted for gestational age. The natural direct, indirect, and total effects of mediation analysis were estimated combining the mediator and the outcome models. Analyses were conducted with and without a miRNA-gestational age interaction term in the model for birthweight. miRNA that showed a significant interaction with gestational age in relation to birthweight, we also estimated the effect of those miRNAs on birthweight at each strata of gestational age, also termed as controlled direct effect in mediation analysis. Direct and indirect effects were averaged across all study participants. Sensitivity analysis was conducted to test the possibility of reverse causation between miRNA expressions and gestational age in relation to birthweight. Bias corrected standard errors were computed from 10000 bootstraps.

Traditionally, the natural direct, indirect and total effects are estimated comparing a unit change in the exposure variable from zero. However, expressions of candidate miRNAs are unlikely to be zero in normal homeostatic condition. Therefore, we scaled our results for a standard deviation change in miRNA expressions from a value that is likely to be the natural placental expression level for that miRNA in a normal pregnancy, i.e. the average placental expression level of a respective miRNA in a 40-week pregnancy, which is the average duration of singleton term pregnancy.

To investigate the role of placenta-derived miRNAs on arsenic-birthweight relation, we first identified miRNAs that are associated with cord blood arsenic by means of multivariate linear regression models adjusting for study site and maternal age. Then, we conducted mediation analysis for the associations between cord blood arsenic and birthweight considering miRNA expressions and gestational age as sequential mediators, where miRNA expressions

temporarily precedes gestational age in the causal pathways (Figure 3.2) following the method described by VanderWeele and Vansteelandt(VanderWeele and Vansteelandt 2014).

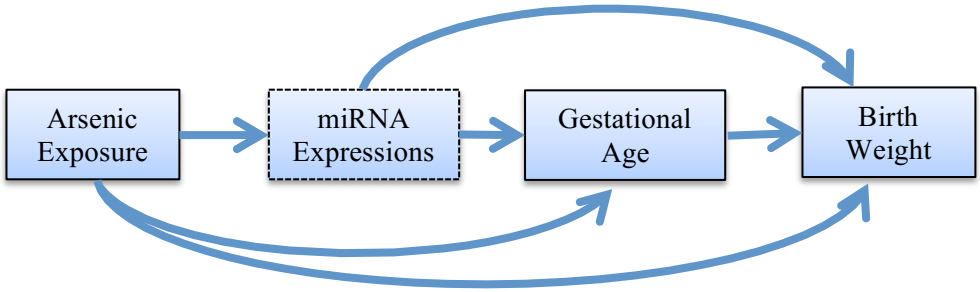

Figure 3.2. Directed acyclic graph showing conceptual model for causal mediation analysis of prenatal arsenic exposure and birthweight considering miRNA expressions and gestational age as sequential mediators

We used multivariate linear regression to model for miRNA expressions and gestational (the mediator models) and for birthweight (the outcome model). Covariates included in the models were natural log cord blood arsenic, study site, maternal age, education, number of past pregnancies, enrollment BMI, and infant gender. The model for gestational age was additionally adjusted for miRNA expressions, while the model of birthweight was additionally adjusted for both miRNA expressions and gestational age. The model of birthweight also included an arsenic-miRNA interaction term, although a miRNA-gestational age interaction term could not be accommodated due to methodological limitations. The outcome and mediator models were combined to estimate the pathway-specific effects of arsenic on birthweight.

Mediation analysis assumes that conditional on the covariates, there is no confounding of the exposure-outcome relation, exposure-mediator relation, mediator-outcome relation and that there is no effect of the exposure that itself confounds the mediator-outcome relation(Valeri and Vanderweele 2013). In all analyses, R version 3.2.2 (Austria, Vienna) or STATA version 14.0 (College Station, TX), 2-sided tests, and significance level of 0.05 were used.

All protocols were reviewed and approved by the Human Research Committees at Harvard T.H. Chan School of Public Health and Dhaka Community Hospital Trust. All participants provided written informed consent prior to participation in the study.

Results

The characteristics of the 77 cases and 77 controls from the discovery study as well as participants from the replication and full cohort were summarized in Table 1. Infant birthweight, gestational age, and maternal enrollment BMI, education, study site, and cord blood arsenic concentrations significantly differed between the case and control groups ($P<0.05$). The mean gestational age among cases was 35.1 weeks (range: 28-41) compared to 38.5 weeks (range: 36-41) among controls. 87% of the cases originated from the Pabna cohort compared to 57.1% of the controls. Cases also showed higher median exposure in cord blood arsenic (i.e. 0.7 μg/dl) compared to the controls (i.e. 0.54 μg/dl). However, the replication cohort (n=364) selected randomly from the whole study cohort (n=1,142) appeared to be similar in key demographic characteristics, birth measurements and exposure levels. The miRNA assay quality results between the discovery and replication cohort appear in the Supplemental Table 3.2.

Discovery analysis results

Associations between miRNA expressions and gestational age: Of 306 miRNAs that passed quality controls in the discovery study; multivariate linear regression adjusted for study site, and maternal age demonstrated 24 miRNAs significantly associated with gestational age at permutation *P*-value<0.01 in at least two out of four normalization methods (Supplemental Table 3.3). Sixteen miRNAs (i.e. let-7g, miR-138, miR-489, miR-542-3p, miR-16, miR-34b, miR-199a, miR-181a, miR-296, miR-205, miR-328, miR-195, miR-886-3p, miR-27a, miR-331, and

miR-181c) passed the significance threshold in three or more normalization methods. miR-1290 showed marginal significance (P=0.01) in one normalization method in the pooled analysis, but showed strong association among cases, and was subsequently selected for further analysis. Among the 25 miRNAs selected for replication, 10 miRNAs (miR-138, miR-489, miR-629, miR-16, miR-34b, miR-522, miR-1290, miR-195, miR-324-5p, and miR337-3p) were found to be negatively associated with gestational age, and the rest 15 were positively associated with gestational age.

Table 3.1. **Distributions of selected characteristics of the study participants by study cohort**

Characteristics	**Discovery Cohort**			**Replication Cohort**	**Full Cohort**
	Control (n=77)	Case (n=77)	P-value	(n=364)	(n=1,142)
Birth weight, (kg)	3.5 (3.3−3.9)	2.0 (0.8−2.2)	<0.0001	2.8 (1.02−4.25)	2.8 (0.8−4.8)
Maternal Age (years)	23.5 (18−38)	23.3 (18−40)	0.7	22.7 (18−38)	23.0 (18−41)
Gestational age (weeks)	38.5 (36−41)	35.1 (28−41)	<0.0001	37.9 (29−41)	37.9 (22−42)
Enrollment BMI, n (%)					
≤18·5 kg/m^2	12 (15.6)	23 (29.9)	0.02	92 (25.3)	315 (27.8)
>18·5 to ≤25.0 kg/m^2	53 (68.8)	50 (64.9)		239 (65.7)	712 (62.8)
>25.0 kg/m^2	12 (15.6)	4 (5.2)		33 (9.0)	107 (9.4)
Maternal Education, n (%)					
No formal education	5 (6.5)	18 (23.4)	0.01	57 (25.7)	162 (14.3)
Primary education	21 (27.3)	20 (26.0)		127 (34.9)	370 (32.6)
Secondary or higher	51 (66.2)	39 (50.6)		180 (49.4)	602 (53.1)
Infant sex, n (%)					
Female	43 (55.8)	38 (49.4)	0.4	179 (49.2)	578 (50.7)
Male	34 (44.2)	39 (50.6)		185 (50.8)	563 (49.3)
Study site, n (%)					
Sirajdikhan	33 (42.9)	10 (13.0)	<0.001	173 (47.5)	575 (50.4)
Pabna	44 (57.1)	67 (87.0)		191 (52.5)	567 (49.6)
No. of past pregnancy					
0	27 (35.1)	26 (33.8)	0.5	143 (39.5)	460 (40.3)
1	28 (36.4)	23 (29.9)		118 (32.6)	339 (29.7)
≥ 2	22 (28.6)	28 (36.4)		101 (27.9)	343 (30.0)
Cord blood arsenic, (μg/dl)	0.54 (0.1-7.1)	0.70 (0.09-6.0)	0.003	0.60 (0.1-15.5)	0.54 (0.1−23.4)

*Values are mean (range) except for cord blood arsenic exposure, which are median (range)

Associations between miRNA expressions and the case-control status: Our analyses revealed 33 miRNAs that were differentially expressed between cases and controls in at least one normalization method; 23 miRNAs in two or more normalization methods at the level of permutation *P*-value<0.01 (Supplemental Table 3.4). Thirteen miRNAs passed the significance threshold in three or more normalization methods (i.e. miR-124a, miR-625, miR-524, miR-518f, miR-376b, miR-34c, miR-337-3p, miR-324-5p, miR-30c, miR-205, miR-195, miR-185).

Total 49 miRNAs were selected for further analysis in the replication cohort based their associations with the case control status (n=24), or gestational age (n=16) or both (n=9) (Table 3.2).

Table 3.2. List of miRNAs selected for replication study

miRNAs associated with case-control status	**miRNAs associated with both case-control status and gestational Age**	**miRNAs associated with gestational age**
1. miR-124a	1) miR-138	1) miR-1290
2. miR-127	2) miR-195	2) miR-16
3. miR-130a	3) miR-205	3) miR-181a
4. miR-135b	4) miR-324-5p	4) miR-181c
5. miR-149	5) miR-337-3p	5) miR-199a
6. miR-15a	6) miR-34b	6) miR-27a
7. miR-185	7) miR-522	7) miR-296
8. miR-200c	8) miR-625.	8) miR-320B
9. miR-20b	9) Let-7g	9) miR-328
10. miR-21		10) miR-331
11. miR-222		11) miR-373
12. miR-30c		12) miR-489
13. miR-34c		13) miR-542.3p
14. miR-376b		14) miR-886.3p
15. miR-424		15) miR-99b
16. miR-454		16) miR-629
17. miR-455		
18. miR-518e		
19. miR-518f		
20. miR-524		
21. miR-720		
22. miR-9		
23. miR-99a		
24. miR-99b.		

Replication analysis results

Associations between miRNA expressions and gestational age: In the replication cohort, 8 miRNAs (e.g. miR-1290, miR-195, miR-328, miR-324-5p, miR-199a, miR-424, let-7g, and miR-27a) showed statistically significant associations with gestational (P<0.05), 4 (e.g. miR-1290, miR-195, miR-328, and miR-324-5p) remained significant after adjusting for multiple testing (Table 3.3). Among the 8 miRNAs associated with gestational age, negative associations were observed for miR-1290, miR-195, let-7g, and miR-27a, while the associations were positive for miR-328, miR-324-5p, miR-199a, and miR-424. Stronger associations in terms of both magnitude and statistical significance were observed for miR-1290, miR195, miR-328 and miR-324-5p, where each standard deviation increase in expression were associated with -0.7 weeks (95% CI: -0.9, -0.4), -0.6 weeks (95% CI: -0.8, -0.3), 0.4 weeks (95% CI: 0.2, 0.6), and 0.4 weeks (95% CI: 0.1, 0.7) changes in gestational age, respectively.

Associations between miRNA expressions and birthweight: Mediation analysis identified the same 8 miRNAs associated with gestational age also associated with birthweight via indirect pathways (P<0.05). Four miRNAs: miR-1290, miR-489, miR-138, and miR-720 were also associated with birthweight via pathways independent of gestational age, while miR-1290, miR-195, miR-489, miR-138, and miR-720 were associated with birthweight via both pathways (Table 3.3). After accounting for multiple testing, the indirect effect associations between miRNA expressions and birthweight remained significant for miR-1290, miR-195, miR-328, and miR-324-5p (FDR q<0.05). Two of these miRNAs: miR-1290 miR-195 was negatively associated with birthweight, while miR-328 and miR-324-5p was positively associated with birthweight.

In terms of the magnitude of association, each standard deviation increase in miR-1290 expression was associated with 136.9g (95% CI: -213.7, -49.9) decreased birthweight; 41.4% of which was mediated via decreasing gestational age (-56.5g, 95% CI: -100.6, -24.3), and the rest (-80.3g, 95% CI: -146.2, -6.5) was operated via pathways independent of gestational age. Similarly, each standard deviation increase in miR-195 expression was associated with 89.4g (95% CI: -175.9, -1.5) decreased birthweight; 43.3% of which was mediated through decreasing gestational age (-42.8g; 95%CI: -83.2, -15.1), and the rest (-46.6g; 95% CI: -123.3, 34.7) was operated via pathways independent of gestational age.

While the associations between miRNAs expressions and birthweight via indirect and direct effect pathways were in the same direction for miRNA-1290, miR-195, miR-328, miR-324-5p, the associations were opposite for miR-199a, miR-27a, and miR-424. For instance, each standard deviation increase miR-199a expression was associated with 36.4g (95% CI: 8.1, 72.8) increased birthweight via increasing gestational age, while 68.5g (95%CI: -161.0, 10.6) decreased birthweight via pathways independent of gestational age, although later association was not statistically significant. In addition, while miR-489 and miR-138 did not show significant associations with birthweight either via indirect or direct pathways, their overall effect on birthweight via both pathways were significant ($P<0.05$).

The effect of miRNA on birthweight varies by week of pregnancy: Our analyses also revealed that the effect of miR-1290, miR-195, miR-328, and miR-489 on birthweight were different by gestational age strata. For instance, the negative associations between miR-1290, miR-195, and miR-489 and birthweight as well as the positive associations between miR-328 and birthweight were stronger in magnitude and statistical significance for pregnancies that terminated early (Figure 3.3, Supplemental Table 3.5). In other words, higher expression of

Table 3.3. Effect of a standard deviation change in placenta-derived miRNA expressions on gestational age [weeks] and birthweight [g]. The total effect of miRNA expressions on birthweight was decomposed into indirect effect mediated through gestational age, and direct effect via pathways independent of gestational age, allowing for miRNA*gestational age interactions in the replication cohort (n=364)

	Associations with Gestational Age			Associations with Birthweight								
				Indirect Effects			Direct Effects			Total Effects		
miRNA	β [g] (95% CI)	*P*-value	FDR	β [g] (95% CI)	*P*-value	FDR	β [g] (95% CI)	*P*-value	FDR	β [g] (95% CI)	*P*-value	FDR
mir1290	-0.7 (-1, -0.4)	<0.001	0.001	-56.5 (-100.6, -24.3)	0.000	0.01	-80.3 (-146.2, -6.5)	0.033	0.46	-136.9 (-213.7, -49.9)	0.002	0.05
mir195	-0.6 (-0.9, -0.3)	<0.001	0.001	-42.8 (-83.2, -15.1)	0.001	0.03	-46.6 (-123.3, 34.7)	0.253	0.73	-89.4 (-175.9, -1.5)	0.046	0.34
mir328	0.4 (0.2, 0.6)	<0.001	0.001	19.9 (6, 43.3)	0.003	0.05	11.3 (-37.7, 57.6)	0.723	0.86	31.2 (-14, 76.3)	0.206	0.77
mir3245p	0.4 (0.1, 0.7)	0.003	0.034	28.3 (9.9, 56.5)	0.004	0.05	-11.1 (-70.7, 45.6)	0.706	0.86	17.2 (-41.9, 74.8)	0.578	0.81
mir489	-0.2 (-0.4, 0)	0.105	0.462	-14.9 (-39.3, 3.7)	0.164	0.60	-61.7 (-112.9, -0.3)	0.044	0.46	-76.6 (-137.5, -7.6)	0.030	0.29
mir199a	0.4 (0.1, 0.8)	0.025	0.156	36.4 (8.1, 72.8)	0.014	0.11	-68.5 (-161, 10.6)	0.118	0.62	-32.1 (-123.6, 57.5)	0.484	0.81
mir27a	-0.4 (-0.8, -0.1)	0.019	0.136	-35.4 (-67.3, -10)	0.016	0.11	31.8 (-69, 97)	0.534	0.85	-3.6 (-115.9, 62.1)	0.894	0.97
mir424	0.3 (0.1, 0.6)	0.010	0.102	26.4 (6.1, 57.1)	0.012	0.11	-19.7 (-81.8, 38.3)	0.600	0.85	6.7 (-53.7, 66.8)	0.812	0.97
let7g	-0.5 (-0.9, -0.1)	0.018	0.136	-40.3 (-89.2, -4.5)	0.030	0.18	10.9 (-83.9, 114.7)	0.892	0.90	-29.4 (-139.8, 84.6)	0.550	0.81
mir138	-0.2 (-0.4, 0.1)	0.203	0.662	-13.1 (-38.8, 4)	0.162	0.60	-48.7 (-100.7, -2.2)	0.044	0.46	-61.8 (-123.5, -10.5)	0.024	0.29
mir720	0.1 (-0.1, 0.3)	0.289	0.708	9.1 (-7.5, 27.3)	0.296	0.73	73.1 (26.2, 116.4)	0.001	0.07	82.2 (30.1, 128.5)	0.002	0.05
mir181c	-0.1 (-0.4, 0.1)	0.369	0.723	-9.6 (-35.4, 9.3)	0.360	0.73	-58.1 (-121.9, 0.2)	0.054	0.46	-67.6 (-133.2, -8.5)	0.024	0.29

*Models were adjusted for maternal age, education, enrollment BMI, no. of past pregnancy, infant gender, study site, and surrogate variables

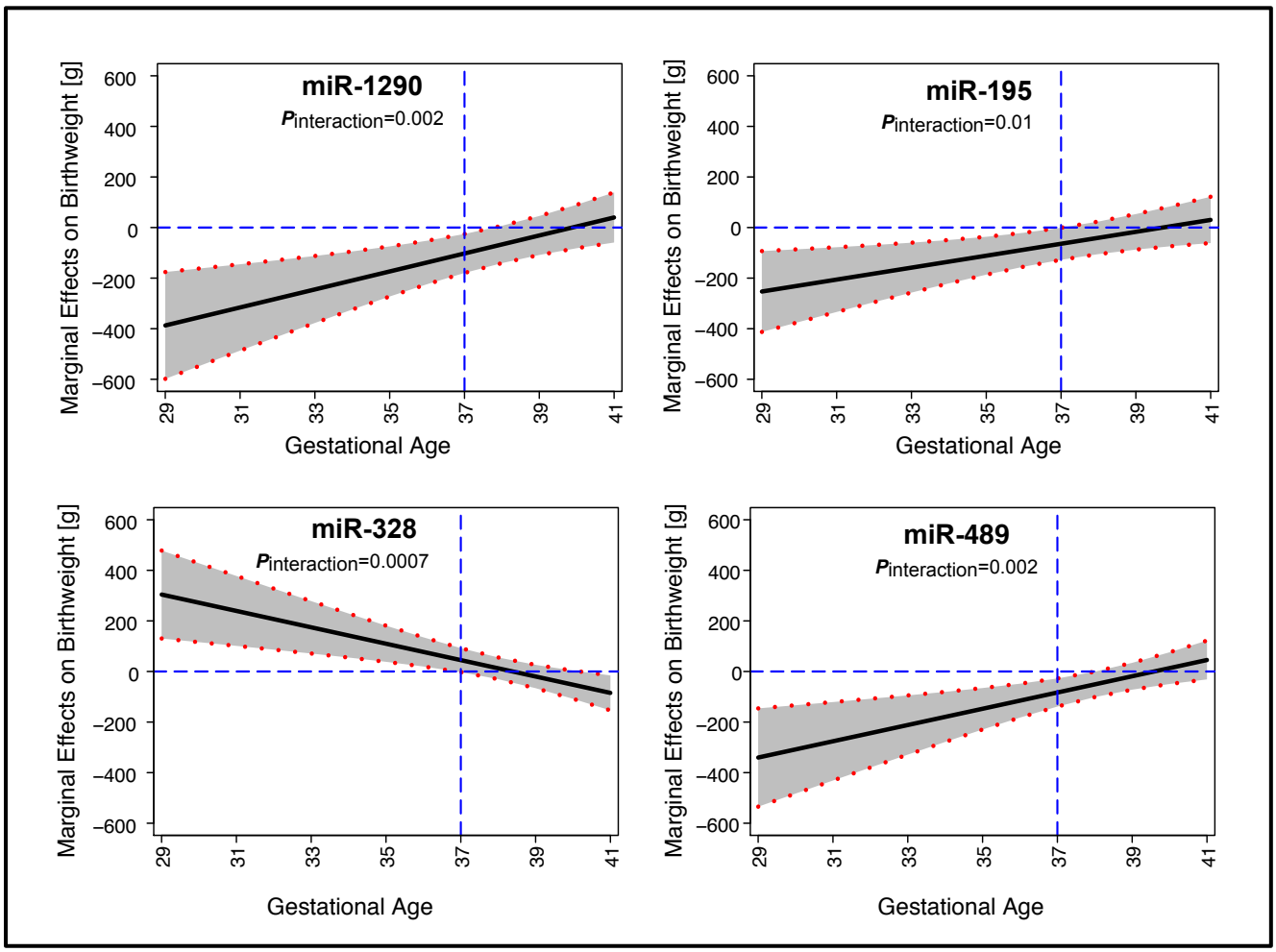

Figure 3.3. Changes in birthweight by a standard deviation change in the expression of selected miRNAs in the placenta by week of pregnancy

miR-1290, miR-195, and miR-489 and lower expression of miR-328 was associated with earlier termination of pregnancy, resulting in lower birthweight.

Associations between cord blood arsenic and miRNA expressions in the placenta: Among 49 miRNAs screened for replication analyses, our results showed that cord blood arsenic exposure was associated with placental expression of 12 miRNAs: miR-518f, miR-524, miR-99b., miR-149, miR-205, miR-328, miR-376b, miR-1290, miR-337-3p, miR-20b, miR-99b, and miR-331 at the significance level of $P<0.05$. Cord blood arsenic exposure was associated with

decreased expression of miR-518f, miR-524, miR-99b., miR-376b, miR-1290, and miR-337-3p and increased expression of miR-149, miR-205, miR-328, miR-20b, miR-99b, and miR-331 at the significance level of $P<0.05$.

Table 3.4. Effects of natural log cord blood arsenic exposure on miRNA expressions in the placenta (n=364)

miRNA	β (95% CI)	P-value	FDR
mir-518f	-0.008 (-0.012, -0.003)	0.001	0.05
mir-524	-0.007 (-0.012, -0.003)	0.002	0.05
mir-99b.	-0.008 (-0.013, -0.002)	0.006	0.08
mir-149	0.006 (0.002, 0.01)	0.006	0.08
mir-337-3p	-0.004 (-0.008, -0.001)	0.023	0.14
mir-376b	-0.008 (-0.015, -0.001)	0.023	0.14
mir-20b	0.005 (0.001, 0.009)	0.025	0.14
mir-205	0.005 (0.001, 0.009)	0.025	0.14
mir-328	0.004 (0.001, 0.008)	0.026	0.14
mir-331	0.003 (0, 0.007)	0.028	0.14
mir-1290	-0.01 (-0.02, -0.001)	0.037	0.16
mir-99b	0.003 (0, 0.006)	0.038	0.16

* Models were adjusted for maternal age, study site, and surrogate variables

The effect of arsenic on birthweight varies by miRNA expression profiles in the placenta:

Our results indicated that the effect of cord blood arsenic exposure on birthweight varies by placental expression of miR-1290 and miR-195 in a dose-dependent manner. Figure 3.4 shows marginal effects of natural log arsenic exposure on birthweight by 1) percentiles of miRNA expressions in the placenta (25^{th}, 50^{th}, and 75^{th}) (blue line plots), and 2) average expression of respective miRNA in preterm (<37 weeks), early term (37-38 weeks), and term (>38 weeks) pregnancies (red line plots). For example, the change in birthweight for a unit increase in natural

log cord blood arsenic exposure was -18.2g (95% CI: -74.6, 38.2) when placental expression of miR-1290 was at 25th percentile, -63.6g (95% CI: -117.0, -10.3) at 50th percentile, and -104.4 (95% CI: -171.0, -37.9) at 75th percentile. Similarly, the effect of the same unit increase in cord blood arsenic exposure on birthweight was -75.2g (95% CI: -130.9, -19.4) when the expression of miR-1290 was set at a level equal to the average expression in preterm pregnancy (<37 weeks), -62.8g (95% CI: -116.0, -9.6) in early-term pregnancy (37-38 weeks), and -50.3g (95% CI: -102.4, 1.8) in term pregnancy (>38 weeks). Similar trend was observed for miR-195.

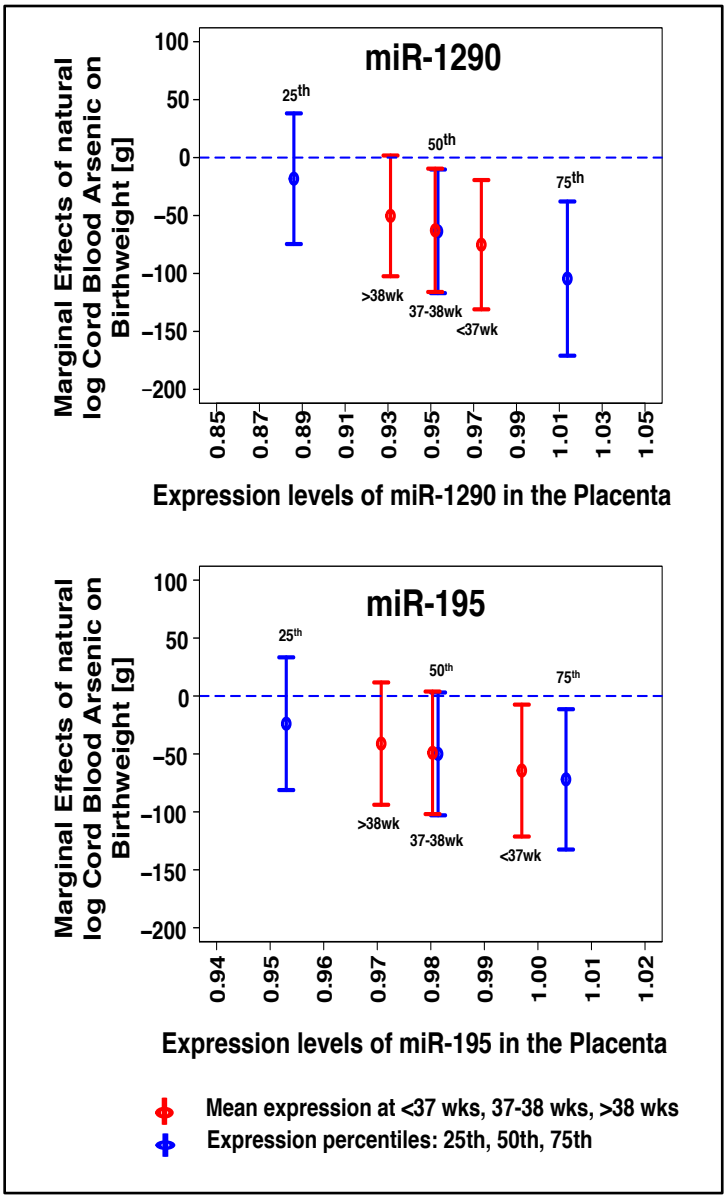

Figure 3.4. Marginal effects of cord blood arsenic on birthweight by miRNA expression levels in the placenta

Placenta-derived miRNAs may offset the harmful effect of arsenic on birthweight: To understand the role of placenta-derived miRNAs on birthweight in response to environmental arsenic exposure, we conducted mediation analysis for the associations between cord blood arsenic and birthweight considering miRNA expressions and gestational age as sequential mediators, where miRNA expressions temporarily precedes gestational age in the causal pathway

(Figure 3.2). Mediation analysis revealed that cord blood arsenic exposure was associated with decreased birthweight, which was primarily mediated via decreasing gestational age (Table 3.5). However, the effects mediated via miRNA expressions in the placenta for miR-1290, miR-195, and miR-328 was operated in the opposite direction, indicating possible feedback mechanism by the placenta to counter harmful effects of arsenic exposure by regulating miRNA expression profiles.

Table 3.5. Direct and Indirect effects of natural log cord blood arsenic exposure on birthweight considering miRNA expressions in the placenta and gestational age as sequential mediators

miRNA	**Effects mediated via miRNA**	**Effects mediated via gestational age**	**Effects mediated via both miRNA & gestational age**	**Effects via pathways other than miRNA & gestational age**
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
miR1290	24.1 (-1, 56.6)	-20.4 (-44.4, -0.5)	3.7 (-25.9, 37.4)	-70.7 (-186.7, 38.1)
miR-195	9.2 (-9.2, 31.9)	-16.5 (-39.6, 2.3)	-7.3 (-34.3, 17.4)	-39 (-138.9, 79.4)
miR-328	10 (-4.2, 31)	-19 (-42.8, 0.2)	-9 (-34.6, 18.2)	-75.3 (-143.9, -19.2)

* Due to methodological limitations, analysis was conducted without a miRNA-gestational age interaction term in the model for birthweight.

Temporal relation between miRNA expressions and gestational age in relation to birthweight: To investigate whether our assumption that miRNA expressions temporarily precedes gestational age in the causal pathways and not the other way around, we conducted sensitivity analysis using three set of models: model for birthweight adjusting for 1) gestational age alone; 2) miRNA alone; and 3) both miRNA and gestational age. Effect estimates of gestational age and miRNA were compared between the models (Table 3.6). Our results showed that the effect estimate of gestational age were largely consistent between model 1 and model 3,

while the effect estimates of miRNAs dramatically changed between model 2 and model 3. This suggested that gestational age was more likely an intermediate or mediating variable between miRNA and birthweight, and that a larger proportion of the total effect of miRNA on birthweight (model 2) was mediated via gestational age (difference between the coefficients of model 2 and model 3).

Table 3.6. Temporal relation between miRNA expressions and gestational age in relation to birthweight

miRNA	**β coefficient for Gestational Age**		**β coefficient for miRNA**	
	Model 1	Model 3	Model 2	Model 3
miR-1290	81.2	76.7	-120.8	-66.4
miR-195	81.2	76.9	-101.4	-55.6
miR-328	81.2	80.3	40.8	7.9
miR-489	81.2	79.4	-60.8	-44.8
miR-324-5p	81.2	81.0	35.7	3.7
miR-720	81.2	79.3	79.7	69.5
miR-199a	81.2	83.7	-34.4	-72.4

Model 1: BW=GA+ C; Model 2: BW=miRNA + C; Model 3: BW= GA + miRNA + C
BW=birthweight, GA=gestational age and C=covariate matrix

In-silico target mapping and miRNA pathway and network analysis: To explore the functional significance of top-hit miRNAs, we predicted mRNA targets associated with miR-1290, miR-195, miR-328, and miR-489 using TaregtScan in MetaCore, GeneGo bioinformatics platform. Predicted miRNA targets with TargetScan total context score <-0.2 were used to identify biological processes associated with top-hit miRNAs. Several biological processes such as cell growth and proliferation, apoptosis, embryonic development, insulin signaling, reproduction and neurohormonal signaling, angiogenesis, and inflammation were found to be associated with predicted mRNA targets of these 4 miRNAs (Figure 3.5).

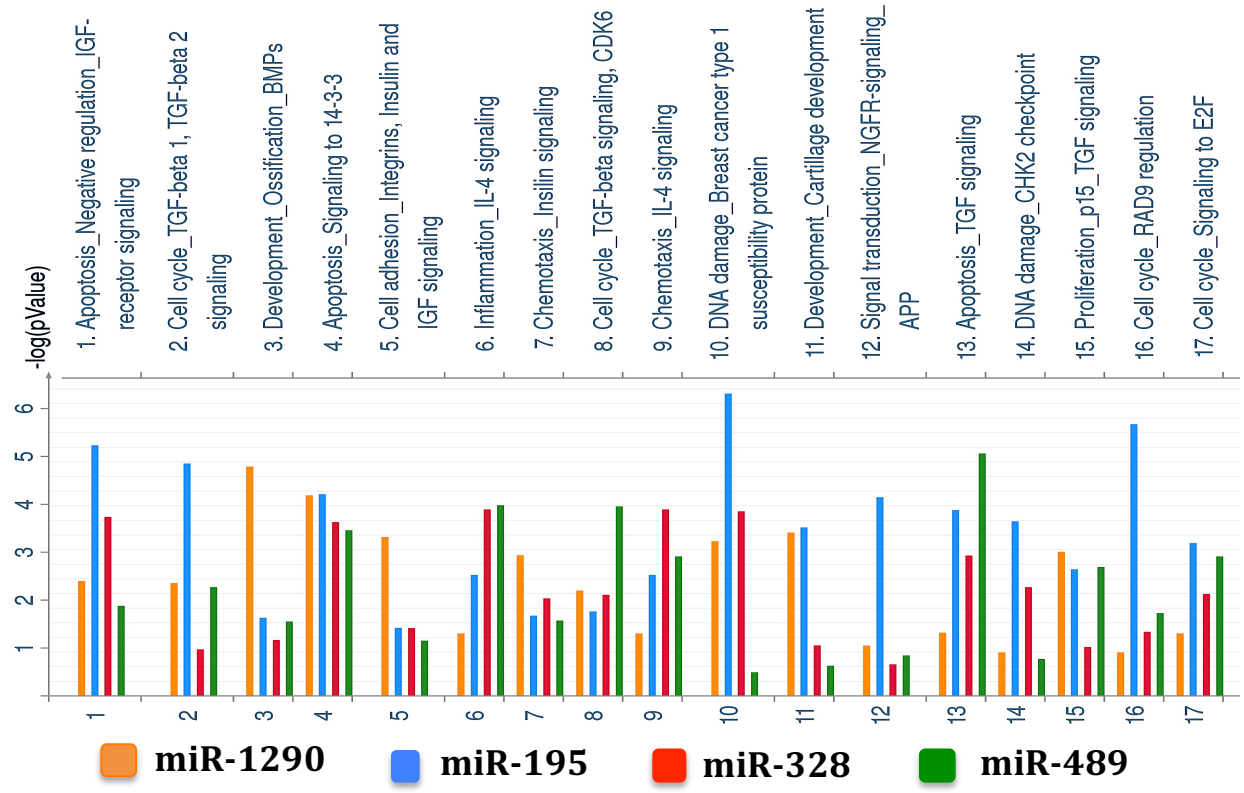

Figure 3.5. Biological Processes associated with predicted mRNA targets of miR-1290, miR-195 and miR-328

Next, we built networks with high trust and curated miRNA targets using MetaCore for miR-1290 and miR-195 (Figure 3.6). Network analysis of miR-195 revealed multiple target genes that are involved in cell growth, proliferation and differentiation and regulation of apoptosis. One of the major target genes of miRNA-195 is the HIF-1 gene, which encodes the alpha subunit of transcription factor hypoxia-inducible factor-1 (HIF-1). HIF-1 functions as a master regulator of cellular and systemic homeostatic response to hypoxia by activating transcription of many genes, including those involved in energy metabolism, angiogenesis, apoptosis, and other genes whose protein products increase oxygen delivery or facilitate

metabolic adaptation to hypoxia. HIF-1 thus plays an essential role in embryonic vascularization, tumor angiogenesis and pathophysiology of ischemic disease. Other major target genes of miR-195 includes FLT3, Bcl-2, M-Ras and A-Raf-1 gene that encodes protein responsible for cell growth and differentiation, apoptosis, proliferation, and differentiation of hematopoietic cells in bone marrow and regulation of apoptosis.

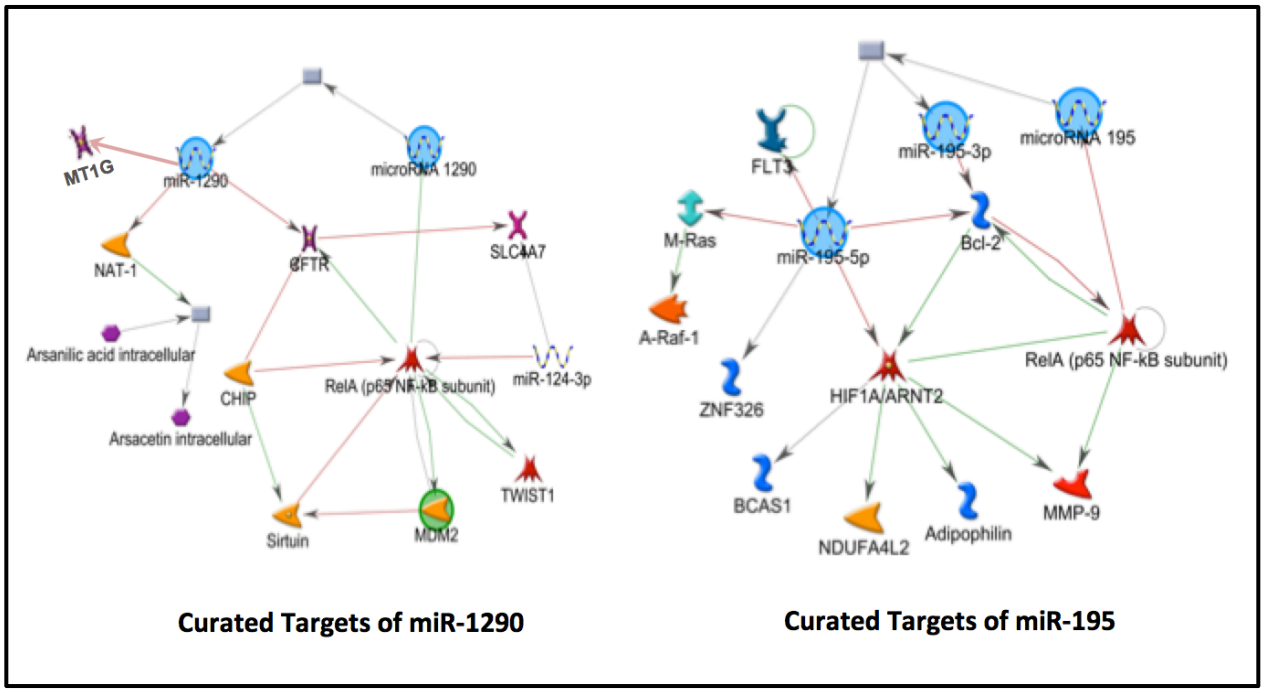

Figure 3.6. Network objects associated with miR-1290 and miR-195

Network analysis of miR-1290 revealed that one of the targets of this miRNA is metallothianine-1G (MT1G) gene. MicroRNA-1290 inhibits MT1G gene, which encodes metallothianine protein, a key transport protein for many metals in the body as well as a potent antioxidant against free radicals.(Ruttkay-Nedecky et al. 2013, Zhang et al. 2016) miR-1290 also inhibits arylamine N-acetyltransferase 1 (NAT-1) gene. The enzyme encoded by this gene catalyzes the transfer of an acetyl group from acetyl-CoA to various arylamine and hydrazine

substrates. This enzyme helps metabolize drugs and other xenobiotics, and functions in folate catabolism(Sim, Abuhammad, and Ryan 2014)

Discussion

To our knowledge, this study is the first of its kind to identify associations between miRNA expression profiles in the placenta and birthweight using an epigenome-wide approach in an attempt to determine the regulation of birthweight by placenta-derived miRNA in response to prenatal arsenic exposure. There has been an increased focus on understanding how prenatal exposure alters the epigenetic trajectories during the critical period of fetal growth and development to alter later life health and disease risk. Our results demonstrated that placenta-derived miRNAs serve as the epigenetic regulator of birthweight by not only determining the length of gestation but also modulating the susceptibility of environmental toxicants on fetal growth

Our results showed that placental expression of several miRNAs were positively or negatively associated with birthweight via pathways that operate through changing gestational age, or pathways that is independent of gestational age, or via both pathways. The effect of miR-1290, miR-195, miR-328, and miR-489 on birthweight varied by gestational age, where the associations were stronger in magnitude and statistical significance in shorter pregnancies, indicating that higher expression of miR-1290, miR-195, and miR-489 and lower expression of miR-328 terminates pregnancy earlier than its natural duration. Our results also showed that placental expression of miR-1290 and miR-195 modulated the negative associations between prenatal arsenic exposure and birthweight, where the susceptibility of the toxic effect of arsenic was stronger for higher expressions of these miRNAs. Together, our results provide molecular basis to support the growing evidence suggesting that smaller infants, who are already at higher

risk of perinatal mortality and morbidity, might be disproportionately affected by environmental toxicants during pregnancy.

We input these three miRNAs into our *in silico* analyses to investigate potential targets and affected pathways. Gene enrichment analysis revealed several biological processes associated with the potential mRNA targets of these three miRNAs. These processes included insulin signaling, regulation of apoptosis, cell growth and proliferation, embryonic development, ossification, reproduction and neurohormonal signaling, angiogenesis, and inflammation. More specifically, network analysis of miR-1290 revealed several curated network objects associated with this miRNA including the MT1G, and NAT-1 genes, where miR-1290 represses the expression of protein products encoded by these genes. NAT-1 gene encodes for several proteins responsible for the metabolism of xenobiotics as well as folate catabolism (Sim, Abuhammad, and Ryan 2014). MT1G gene, on the other hand, encodes metallothianine protein, a key element for metal transport and elimination from human body (Zhang et al. 2016), and also responsible for protecting cells from free radical induced oxidative damage (Ruttkay-Nedecky et al. 2013) Among several mechanisms how inorganic arsenic exerts its adverse health effects, oxidative stress response is most widely studied.(Hubaux et al. 2013) Our results suggest that reduced expression miR-1290 in placental tissues will promotes overexpression of MT1G and NAT-1 gene products to facilitate the transport, metabolism, and elimination of arsenic from the body. On the other hand, miR-195 showed binding sites on several genes including the HIF-1 gene, which encodes the alpha subunit of transcription factor hypoxia-inducible factor-1 (HIF-1).(Gonsalves et al. 2015) HIF-1 functions as a master regulator of cellular and systemic homeostatic response to hypoxia by activating transcription of many genes, including those involved in energy metabolism, angiogenesis, apoptosis, and other genes whose protein products

increase oxygen delivery or facilitate metabolic adaptation to hypoxia.

The association between miRNAs expression profiles and birth outcome was reported in few other studies; although most of those studies used candidate miRNA expression analysis approach and small sample size. For instance, Maccani et al identified that placental expression of miR-16, and miR-21 was associated with poor fetal growth and maternal smoking during pregnancy.(Maccani, Padbury, and Marsit 2011, Maccani et al. 2010) Another study identified placental expression of miR-424, which is an epigenetic regulator of hypoxia, was also associated with poor fetal growth.(Huang et al. 2013) Additionally, few other studies looked at the association between miRNA expression profiles and pregnancy conditions reporting that the expression of miR-17,(Chen and Wang 2013) miR-206,(Akehurst et al. 2015) miR-34a,(Doridot et al. 2014) miR-210 and miR-34c-5p (Enquobahrie et al. 2011), miR-126,(Hong, Li, and Xu 2014) were associated with preeclampsia, which is one of the risk factors for poor fetal growth. On the other hand, very few studies investigated whether in-utero exposure to inorganic arsenic is associated with altered expression of miRNAs in the fetal tissues. For instance, Rager et. al identified 12 miRNAs associated with maternal total urinary arsenic collected at the time of delivery that were associated with immune signaling and inflammation. However, ours is the first study to link between in-utero arsenic exposure and birthweight via miRNA expression profiles in placenta using global miRNA expression analysis approach.

Our study has some limitations. Our replication cohort was not completely independent of the discovery cohort as 52 samples from the discovery cohort were also included in the replication analysis. Because we employed extreme phenotype sampling to select cases and controls from both tails of the birthweight distribution in the discovery cohort, selecting a representative sample of the full cohort protected us from penalizing for the lost in variance in

birthweight distribution in the replication analysis. Future studies designed to replicate our discovery findings in an external population will be very useful. Additionally, our discovery analysis platform for miRNA expression analysis using TaqMan qRT-PCR miRNA Array lacked epigenome-wide coverage and was designed for 754 preconfigured human miRNAs based on miRBase 14. To date, nearly 3,000 mature human miRNAs has been discovered and many more probably are yet to be discovered. Therefore, we might have missed some of the miRNAs, including some novel miRNAs that could be potentially associated with our exposure and outcome of interests.

Our study has several strengths. The two-stage study design allowed us to validate the expression of top-hit miRNAs in the replication cohort using TaqMan qRT-PCR technology, which is considered as gold standard for miRNA expression analysis and is usually recommended to confirm results from microarray data. However, in two-stage study designs, technical variations may arise between expression data of the discovery and replication phase if different expression technology is used, which was not in our case. The miRNA expression analysis platforms we employed in the discovery and replication phase both are based on qRT-PCR technology, reducing the chance of colluding our study findings due to instrumental variations. Moreover, in both phases, we employed surrogate variable analysis in order to adjust for known or unknown biological or technical variations, which are common in gene expression data. (Leek and Storey 2007)

Epigenetic signals including miRNA expression are tissue-sensitive and therefore, it is important to assess miRNA expressions in tissues that are biologically relevant with the exposure and outcome of interest. The placenta is the principal organ to maintain fetal growth and development by regulating the intrauterine environment the growing fetus is exposes to. Thus,

the placenta is likely to express the strongest signals of any perturbations in the intrauterine environment as well as molecular markers that might be associated with birthweight. However, as is the case for all biological specimens with a heterogeneous cell type composition, the placental miRNA content can vary based on the sampling method implemented. In the current study, consistency was maintained by removing the maternal decidua prior to collecting placental tissue samples the fetal side of the placenta.

In our study, we showed that miRNA expressions in the placenta temporally precede gestational age, which validates the temporality assumption of our causal mediation modeling approach. Our use of cord blood samples to characterize arsenic exposure allowed us to have the most accurate measure of fetal exposure to arsenic *in-utero*. The prospective nature of this study with fairly a good number of sample sizes is another strength. Finally, by identifying miRNA expression markers in the placenta associated with both birthweight and *in-utero* exposure to environmental toxicants during the critical period of fetal development, this research provides evidence that alteration of epigenetic signaling is a critical mode of the regulatory function of the placenta to maintain the growth and development of the growing fetus.

Conclusions

Our study provides evidence indicating that placenta-derived miRNAs regulate birthweight by not only determining the length of gestation but also modulating the susceptibility of environmental toxicants on fetal growth. It will be useful to validate the findings from this study in an external cohort.

Bibliography

Adibi, J. J., R. Hauser, P. L. Williams, R. M. Whyatt, H. M. Thaker, H. Nelson, R. Herrick, and H. K. Bhat. 2009. "Placental biomarkers of phthalate effects on mRNA transcription: application in epidemiologic research." *Environ Health* 8:20. doi: 10.1186/1476-069X-8-20.

Akehurst, C., H. Y. Small, L. Sharafetdinova, R. Forrest, W. Beattie, C. E. Brown, S. W. Robinson, J. D. McClure, L. M. Work, D. M. Carty, M. W. McBride, D. J. Freeman, and C. Delles. 2015. "Differential expression of microRNA-206 and its target genes in preeclampsia." *J Hypertens* 33 (10):2068-74. doi: 10.1097/hjh.0000000000000656.

Akey, J. M., S. Biswas, J. T. Leek, and J. D. Storey. 2007. "On the design and analysis of gene expression studies in human populations." *Nat Genet* 39 (7):807-8; author reply 808-9. doi: 10.1038/ng0707-807.

Barker, D. J. 2004. "The developmental origins of chronic adult disease." *Acta Paediatr Suppl* 93 (446):26-33.

Bentwich, I., A. Avniel, Y. Karov, R. Aharonov, S. Gilad, O. Barad, A. Barzilai, P. Einat, U. Einav, E. Meiri, E. Sharon, Y. Spector, and Z. Bentwich. 2005. "Identification of hundreds of conserved and nonconserved human microRNAs." *Nat Genet* 37 (7):766-70. doi: 10.1038/ng1590.

Calin, G. A., and C. M. Croce. 2006. "MicroRNA signatures in human cancers." *Nat Rev Cancer* 6 (11):857-66. doi: 10.1038/nrc1997.

Chen, D. B., and W. Wang. 2013. "Human placental microRNAs and preeclampsia." *Biol Reprod* 88 (5):130. doi: 10.1095/biolreprod.113.107805.

Concha, G., B. Nermell, and M. V. Vahter. 1998. "Metabolism of inorganic arsenic in children with chronic high arsenic exposure in northern Argentina." *Environ Health Perspect* 106 (6):355-9.

D'Haene, B., P. Mestdagh, J. Hellemans, and J. Vandesompele. 2012. "miRNA expression profiling: from reference genes to global mean normalization." *Methods Mol Biol* 822:261-72. doi: 10.1007/978-1-61779-427-8_18.

Doridot, L., D. Houry, H. Gaillard, S. T. Chelbi, S. Barbaux, and D. Vaiman. 2014. "miR-34a expression, epigenetic regulation, and function in human placental diseases." *Epigenetics* 9 (1):142-51. doi: 10.4161/epi.26196.

Enquobahrie, D. A., D. F. Abetew, T. K. Sorensen, D. Willoughby, K. Chidambaram, and M. A. Williams. 2011. "Placental microRNA expression in pregnancies complicated by preeclampsia." *Am J Obstet Gynecol* 204 (2):178.e12-21. doi: 10.1016/j.ajog.2010.09.004.

Gonsalves, C. S., C. Li, M. S. Mpollo, V. Pullarkat, P. Malik, S. M. Tahara, and V. K. Kalra. 2015. "Erythropoietin-mediated expression of placenta growth factor is regulated via

activation of hypoxia-inducible factor-1alpha and post-transcriptionally by miR-214 in sickle cell disease." *Biochem J* 468 (3):409-23. doi: 10.1042/bj20141138.

Gude, N. M., C. T. Roberts, B. Kalionis, and R. G. King. 2004. "Growth and function of the normal human placenta." *Thromb Res* 114 (5-6):397-407. doi: 10.1016/j.thromres.2004.06.038.

Hagen, J. W., and E. C. Lai. 2008. "microRNA control of cell-cell signaling during development and disease." *Cell Cycle* 7 (15):2327-32.

Hong, F., Y. Li, and Y. Xu. 2014. "Decreased placental miR-126 expression and vascular endothelial growth factor levels in patients with pre-eclampsia." *J Int Med Res* 42 (6):1243-51. doi: 10.1177/0300060514540627.

Huang, L., Z. Shen, Q. Xu, X. Huang, Q. Chen, and D. Li. 2013. "Increased levels of microRNA-424 are associated with the pathogenesis of fetal growth restriction." *Placenta* 34 (7):624-7. doi: 10.1016/j.placenta.2013.04.009.

Hubaux, R., D. D. Becker-Santos, K. S. Enfield, D. Rowbotham, S. Lam, W. L. Lam, and V. D. Martinez. 2013. "Molecular features in arsenic-induced lung tumors." *Mol Cancer* 12:20. doi: 10.1186/1476-4598-12-20.

Kile, M. L., E. G. Rodrigues, M. Mazumdar, C. B. Dobson, N. Diao, M. Golam, Q. Quamruzzaman, M. Rahman, and D. C. Christiani. 2014. "A prospective cohort study of the association between drinking water arsenic exposure and self-reported maternal health symptoms during pregnancy in Bangladesh." *Environ Health* 13 (1):29. doi: 10.1186/1476-069X-13-29.

Lee, R. C., R. L. Feinbaum, and V. Ambros. 1993. "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." *Cell* 75 (5):843-54.

Leek, J. T., and J. D. Storey. 2007. "Capturing heterogeneity in gene expression studies by surrogate variable analysis." *PLoS Genet* 3 (9):1724-35. doi: 10.1371/journal.pgen.0030161.

Maccani, M. A., M. Avissar-Whiting, C. E. Banister, B. McGonnigal, J. F. Padbury, and C. J. Marsit. 2010. "Maternal cigarette smoking during pregnancy is associated with downregulation of miR-16, miR-21, and miR-146a in the placenta." *Epigenetics* 5 (7):583-9.

Maccani, M. A., J. F. Padbury, and C. J. Marsit. 2011. "miR-16 and miR-21 expression in the placenta is associated with fetal growth." *PLoS One* 6 (6):e21210. doi: 10.1371/journal.pone.0021210.

Marsit, C. J. 2015. "Influence of environmental exposure on human epigenetic regulation." *J Exp Biol* 218 (Pt 1):71-9. doi: 10.1242/jeb.106971.

Pineles, B. L., R. Romero, D. Montenegro, A. L. Tarca, Y. M. Han, Y. M. Kim, S. Draghici, J. Espinoza, J. P. Kusanovic, P. Mittal, S. S. Hassan, and C. J. Kim. 2007. "Distinct subsets of

microRNAs are expressed differentially in the human placentas of patients with preeclampsia." *Am J Obstet Gynecol* 196 (3):261 e1-6. doi: 10.1016/j.ajog.2007.01.008.

Rager, J. E., K. A. Bailey, L. Smeester, S. K. Miller, J. S. Parker, J. E. Laine, Z. Drobna, J. Currier, C. Douillet, A. F. Olshan, M. Rubio-Andrade, M. Styblo, G. Garcia-Vargas, and R. C. Fry. 2013. "Prenatal arsenic exposure and the epigenome: Altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood." *Environ Mol Mutagen*. doi: 10.1002/em.21842.

Raghunathan TE, Lepkowski JM, Van Hoewyk J. 2001. "A multivariate technique for multiply imputing missing values using a sequence of regression models. ." *Surv Methodol* (27):85-95.

Ruttkay-Nedecky, B., L. Nejdl, J. Gumulec, O. Zitka, M. Masarik, T. Eckschlager, M. Stiborova, V. Adam, and R. Kizek. 2013. "The role of metallothionein in oxidative stress." *Int J Mol Sci* 14 (3):6044-66. doi: 10.3390/ijms14036044.

Sim, E., A. Abuhammad, and A. Ryan. 2014. "Arylamine N-acetyltransferases: from drug metabolism and pharmacogenetics to drug discovery." *Br J Pharmacol* 171 (11):2705-25. doi: 10.1111/bph.12598.

Valeri, L., and T. J. Vanderweele. 2013. "Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros." *Psychol Methods* 18 (2):137-50. doi: 10.1037/a0031034.

VanderWeele, T. J., and S. Vansteelandt. 2014. "Mediation Analysis with Multiple Mediators." *Epidemiol Method* 2 (1):95-115. doi: 10.1515/em-2012-0010.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman. 2002. "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." *Genome Biol* 3 (7):RESEARCH0034.

Zhang, W. C., T. M. Chin, H. Yang, M. E. Nga, D. P. Lunny, E. K. Lim, L. L. Sun, Y. H. Pang, Y. N. Leow, S. R. Malusay, P. X. Lim, J. Z. Lee, B. J. Tan, N. Shyh-Chang, E. H. Lim, W. T. Lim, D. S. Tan, E. H. Tan, B. C. Tai, R. A. Soo, W. L. Tam, and B. Lim. 2016. "Tumour-initiating cell-specific miR-1246 and miR-1290 expression converge to promote non-small cell lung cancer progression." *Nat Commun* 7:11702. doi: 10.1038/ncomms11702.

Zhu, X. M., T. Han, I. L. Sargent, G. W. Yin, and Y. Q. Yao. 2009. "Differential expression profile of microRNAs in human placentas from preeclamptic pregnancies vs normal pregnancies." *Am J Obstet Gynecol* 200 (6):661 e1-7. doi: 10.1016/j.ajog.2008.12.045.

Appendix

Supplementary Table 3.1 Study design and analysis workflow

DISCOVERY PHASE (n=154)	
Study Design	Extreme Phenotype Sampling: • 77 cases with birthweight (0.8-2.2 kg) • 77 controls with birthweight (3.3-3.9 kg)
Expression platform	TaqMan RT-qPCR OpenArray for 754 preconfigured human miRNAs
Quality Control	309 miRNAs passed quality control criteria: Cycle threshold <30; Cq confidence >1.1;Amp score >0.8; Expression >50% samples
Normalization	1) Endogene control; 2) Quantile; 3) Global; 4) Raw Ct
Association analysis	1) miRNAs associated with the case-control status 2) miRNAs associated with gestational age
MicroRNA selection	Total 49 miRNAs associated with either the case-control status, or gestational age, or both were selected for replication based on- 1) Passed permutation *P*-value <0.01 in ≥2 normalization methods, or 2) Passed permutation *P*-value <0.01 in any normalization method but has biological relevance
REPLICATION PHASE (n=364)	
Study Design	364 participants randomly selected from the full cohort (n=1,141) • 312 new participants; 52 duplicates from discovery cohort
Expression platform	TaqMan RT-qPCR Customized OpenArray Platform • 310 samples in triplicate; 56 samples in duplicate
Quality Control	All miRNAs passed quality control criteria: Cycle threshold <30; Cq confidence >1.1; Amp score >0.8; Expression >50% samples
Normalization	Endogene control
Association analysis	1) Direct, indirect, and total effects of miRNAs on birthweight considering gestational age a mediator 2) Associations between miRNA expressions and cord blood arsenic exposure 3) Direct, indirect, and total effects of arsenic on birthweight considering miRNA and gestational age as sequential mediators 4) miRNA network/pathway analysis

Supplementary Table 3.2. Quality of qRT-PCR data

miRNA	Discovery Study				Replication Study			
	No. of missing	Mean Ct	Sds (σ)	Coefficient of variance*	No. of missing	Mean Ct	Sds (σ)	Coefficient of variance*
let-7g	12	21.72	1.09	0.050	0	17.91	1.32	0.074
miR-127	2	17.97	2.15	0.120	0	16.78	1.28	0.076
miR-1290	52	23.23	2.26	0.097	0	18.04	2.25	0.125
miR-130a	1	17.84	1.68	0.094	0	16.09	1.15	0.072
miR-135b	3	21.91	1.66	0.076	0	18.16	1.40	0.077
miR-138	5	22.63	1.71	0.075	0	20.98	1.57	0.075
miR-149	1	17.46	1.43	0.082	0	17.77	1.24	0.070
miR-15a	17	23.53	1.29	0.055	0	21.70	1.37	0.063
miR-16	2	17.01	1.49	0.088	0	14.27	1.39	0.098
miR-181a	15	17.31	1.55	0.090	0	15.34	1.24	0.081
miR-181c	11	23.75	1.42	0.060	3	20.65	1.53	0.074
miR-185	2	23.33	1.34	0.057	0	20.64	1.66	0.080
miR-195	7	21.55	1.27	0.059	0	18.65	1.34	0.072
miR-199a	8	23.99	1.51	0.063	0	22.23	1.38	0.062
miR-200c	7	15.27	1.53	0.100	0	13.85	1.26	0.091
miR-205	2	20.03	1.46	0.073	0	19.51	1.47	0.075
miR-20b	3	22.79	1.26	0.055	0	19.07	1.43	0.075
miR-21	3	16.01	1.58	0.098	0	14.19	1.45	0.102
miR-222	7	15.58	1.56	0.100	0	13.13	1.40	0.107
miR-27a	1	18.16	1.70	0.094	0	17.00	1.19	0.070
miR-296	10	22.33	1.74	0.078	1	21.53	1.45	0.067
miR-30c	12	14.14	2.24	0.158	0	14.24	1.45	0.102
miR-320B	3	21.69	1.70	0.078	0	18.68	1.24	0.067
miR-324-5p	10	21.39	1.93	0.090	0	19.31	1.15	0.059
miR-328	2	19.15	1.63	0.085	0	19.67	1.16	0.059
miR-331	0	18.12	1.76	0.097	0	18.64	1.34	0.072

* Internal control miRNAs
$ CV= standard deviation (sds) / mean Ct.

Supplementary Table 3.2. (Continued) Quality of qRT-PCR data

miRNA	Discovery Study				Replication Study			
	No. of missing	Mean Ct	Sds (σ)	Coefficient of variance*	No. of missing	Mean Ct	Sds (σ)	Coefficient of variance*
miR-337-3p	16	25.33	1.44	0.057	0	21.67	1.14	0.053
miR-34b	33	23.31	1.90	0.082	0	21.62	1.33	0.061
miR-34c	10	24.19	1.42	0.059	0	24.42	1.56	0.064
miR-373	10	23.66	1.75	0.074	0	21.93	1.58	0.072
miR-376b	65	26.24	1.14	0.043	4	25.36	1.55	0.061
miR-424	18	21.07	1.62	0.077	0	17.84	1.08	0.060
miR-454	54	20.36	1.95	0.096	0	16.86	1.47	0.087
miR-455	11	22.07	1.34	0.061	0	18.95	1.21	0.064
miR-489	33	22.18	1.67	0.075	4	18.55	1.66	0.090
miR-518e	1	17.43	1.37	0.079	0	14.24	1.17	0.082
miR-518f	74	16.41	1.66	0.101	1	15.51	1.19	0.077
miR-522	22	16.45	1.29	0.078	0	14.18	1.22	0.086
miR-524	60	24.26	1.27	0.052	0	18.72	1.02	0.054
miR-542-3p	6	23.95	1.53	0.064	0	22.74	1.04	0.046
miR-625	5	23.93	1.12	0.047	0	17.89	1.34	0.075
miR-629	3	23.91	1.19	0.050	0	21.87	1.27	0.058
miR-720	10	15.25	1.53	0.101	0	14.15	1.69	0.120
miR-886-3p	3	19.71	2.29	0.116	0	18.40	1.57	0.085
miR-9	31	23.39	1.19	0.051	0	22.64	1.29	0.057
miR-99a	6	16.10	1.65	0.103	1	13.62	1.64	0.121
miR-99b	1	16.27	1.74	0.107	0	16.02	1.11	0.069
miR-124a	26	24.05	2.53	0.105	0	23.35	1.67	0.072
miR-362*	5	22.90	1.08	0.047	0	19.84	1.26	0.064
miR-532*	1	19.29	1.28	0.067	0	17.75	1.46	0.082
miR-532-3p*	2	20.24	1.37	0.067	0	19.57	1.15	0.059

* Internal control miRNAs
$ CV= standard deviation (sds) / mean Ct.

Supplementary Table 3.3. Selection of miRNA from discovery cohort (n=154) based on their associations with gestational age adjusting for maternal age, infant gender, and study site

miRNA	Endogene		Global		Quantile		Raw Ct	
	β	*P*-value	β	*P*-value	β	P-value	β	P-value
let.7g	-0.74	0.001	-0.57	<0.001	-0.54	0.007	-0.39	0.051
miR.1290	-0.43	0.01	-0.24	0.15	-0.38	0.02	-0.32	0.07
miR.135b	-0.33	0.008	-0.49	0.118	-0.49	0.14	0.19	0.588
miR.138	-1.15	<0.001	-1.02	<0.001	-1.13	<0.001	0.48	0.157
miR.16	-0.68	0.001	-0.77	<0.001	-0.65	0.001	-0.23	0.33
miR.181a	0.51	0.007	0.37	0.03	0.51	0.004	0.5	0.005
miR.181c	0.32	0.08	0.48	0.009	0.55	0.004	0.61	<0.001
miR.195	-1.02	0.009	-0.92	<0.001	-0.73	0.008	-0.39	0.435
miR.199a	0.53	0.004	0.54	<0.001	0.71	<0.001	0.77	<0.001
miR.205	0.52	0.008	0.5	0.005	0.58	0.005	0.71	<0.001
miR.27a	0.44	0.02	0.53	0.003	0.68	<0.001	0.69	0.002
miR.296	0.53	0.007	0.58	0.001	0.7	<0.001	0.72	<0.001
miR.320B	1.0	<0.001	0.56	0.058	0.68	0.055	-0.83	0.089
miR.324.5p	1.31	<0.001	2.24	0.008	1.85	0.014	-0.42	0.73
miR.328	0.5	0.01	0.49	0.005	0.57	0.003	0.67	0.002
miR.331	0.44	0.023	0.51	0.007	0.62	0.001	0.76	<0.001
miR.337.3p	0.62	0.017	0.8	0.001	0.81	0.002	-0.05	0.902
miR.34b	-0.52	0.002	-0.32	0.038	-0.46	0.003	-0.39	0.008
miR.373	0.33	0.068	0.37	0.045	0.53	0.004	0.54	0.003
miR.489	-0.6	<0.001	-0.48	0.009	-0.57	0.003	-0.3	0.102
miR.522	-0.59	0.002	-0.51	0.005	-0.46	0.013	-0.43	0.051
miR.542.3p	0.58	0.001	0.65	<0.001	0.72	<0.001	0.76	<0.001
miR.625.	-0.8	0.002	-0.72	0.105	-1.09	0.002	0.25	0.549
miR.629	-0.67	0.001	-0.39	0.037	-0.65	<0.001	-0.21	0.279
miR.886.3p	0.47	0.015	0.52	0.004	0.54	0.004	0.69	<0.001
miR.99b	0.48	0.012	0.48	0.014	0.72	<0.001	0.7	0.004

*All *P*-values are permutation based *P*-values calculated based on 1000 permutations
*Effect estimates indicate change in gestational age in weeks by a standard deviation change in miRNA expression

Supplementary Table 3.4. Selection of miRNA from discovery cohort (n=154) based on their associations with the case-control status. Model 1 was adjusted for maternal age, infant gender, and study site, while Model 2 was additionally adjusted for gestational age

miRNA	Model 1				Model 2			
	Endogene P-value	Global P-value	Quantile P-value	Raw P-value	Endogene P-value	Global P-value	Quantile P-value	Raw P-value
Let-7g	0.001	0.002	0.032	0.314				
miR.124a	0.084	0.235	0.135	0.713	0.001	0.003	0.002	0.817
miR.127					0.003	0.021	0.002	0.884
miR.130a	0.121	0.041	0.005	0.874				
miR.135b	0.047	0.007	0.032	0.89				
miR.138	0.007	0.008	0.013	0.829				
miR.149	0.082	0.001	0.006	0.642	0.341	0.009	0.096	0.402
miR.15a	0.005	0.166	0.07	0.341	0.002	0.048	0.015	0.488
miR.185	0.006	0.003	0.004	0.826	0.017	0.02	0.001	0.785
miR.195	<0.001	0.001	0.001	0.563	0.012	0.184	0.056	0.462
miR.200c	0.058	0.005	0.018	0.574	0.065	0.001	0.016	0.374
miR.205	0.004	0.002	0.001	0.429				
miR.20b	0.009	0.484	0.33	0.064	0.033	0.014	0.063	0.198
miR.21					0.017	0.239	0.003	0.666
miR.222	0.171	0.002	0.004	0.726	0.128	0.001	0.042	0.215
miR.30c	0.001	0.005	<0.001	0.257	<0.001	0.001	<0.001	0.23
miR.324.5p					0.004	0.007	0.001	0.861
miR.337.3p	<0.001	0.01	0.001	0.203	0.009	0.007	0.006	0.329
miR.34b	<0.001	<0.001	<0.001	0.003	0.068	0.038	0.109	0.002
miR.34c	0.002	0.003	0.001	0.401	0.005	0.083	0.032	0.495
miR.376b	<0.001	0.003	<0.001	0.447	<0.001	<0.001	<0.001	0.804
miR.424	0.001	0.057	0.01	0.889	<0.001	0.013	0.006	0.892
miR.454	0.082	0.005	0.001	0.086	0.006	0.001	<0.001	0.054
miR.455					0.008	0.116	0.102	0.546
miR.518e	0.2358	0.004	0.1049	0.3137	0.022	0.001	0.03	0.279
miR.518f	0.046	<0.001	<0.001	0.003	0.007	<0.001	<0.001	0.004
miR.522					0.017	<0.001	0.001	0.253
miR.524	0.007	0.003	0.002	0.273	0.044	0.008	<0.001	0.001
miR.625.					0.002	0.006	0.001	0.062
miR.629					0.031	0.021	0.003	0.044
miR.9	0.051	0.001	0.001	0.146	0.116	0.006	0.02	0.163
miR.99a	0.03	0.009	0.038	0.905				
miR.99b.	0.383	0.397	0.004	0.539	0.084	0.087	0.009	0.097

* All *P*-values are permutation based *P*-values calculated based on 1000 permutations

Supplementary Table 3.5. . Effect of a standard deviation change in selected placenta-derived miRNA expressions on birthweight [g] by gestational age in the replication cohort (n=364)

Gestational Age	miR-1290		miR-195		miR-328		miR-489	
	β (95% CI)	*P*-value	β (95% CI)	*P*-value	β (95% CI)	*P*-value	β (95% CI)	*P*-value
At 29 wks	-387.1 (-599.2, -175.1)	0.0004	-253 (-490.4, -15.6)	0.002	304.3 (129.8, 478.8)	0.0007	-340.3 (-535.3, -145.2)	0.0007
At 30 wks	-351.6 (-543.1, -160.1)	0.0004	-229.4 (-442.8, -16)	0.002	271.9 (115.4, 428.4)	0.0007	-308.1 (-483.4, -132.9)	0.0006
At 31 wks	-316 (-487.3, -144.6)	0.0003	-205.8 (-395.9, -15.7)	0.002	239.5 (100.8, 378.2)	0.0008	-276 (-431.6, -120.3)	0.0006
At 32 wks	-280.4 (-432.2, -128.7)	0.0003	-182.2 (-349.8, -14.6)	0.002	207.1 (85.9, 328.2)	0.0009	-243.8 (-380.2, -107.5)	0.0005
At 33 wks	-244.8 (-377.8, -111.9)	0.0003	-158.6 (-305.1, -12.1)	0.002	174.6 (70.7, 278.6)	0.001	-211.7 (-329.2, -94.2)	0.0004
At 34 wks	-209.3 (-324.5, -94)	0.0004	-135 (-262.3, -7.7)	0.002	142.2 (54.9, 229.5)	0.001	-179.6 (-278.9, -80.2)	0.0004
At 35 wks	-173.7 (-273, -74.3)	0.0007	-111.4 (-222.6, -0.2)	0.004	109.8 (38.1, 181.5)	0.003	-147.4 (-229.7, -65.2)	0.0005
At 36 wks	-138.1 (-224.3, -51.9)	0.002	-87.8 (-187.3, 11.7)	0.01	77.4 (19.5, 135.3)	0.009	-115.3 (-182.5, -48.1)	0.0008
At 37 wks	-102.5 (-179.8, -25.2)	0.009	-64.2 (-158.2, 29.7)	0.05	45 (-2.5, 92.5)	0.06	-83.1 (-138.9, -27.4)	0.004
At 38 wks	-67 (-141, 7.1)	0.08	-40.6 (-136.3, 55)	0.22	12.5 (-30.5, 55.6)	0.57	-51 (-101.5, -0.6)	0.05
At 39 wks	-31.4 (-108.7, 45.9)	0.43	-17 (-121.2, 87.2)	0.64	-19.9 (-66.2, 26.4)	0.40	-18.9 (-72.1, 34.3)	0.49
At 40 wks	4.2 (-82.1, 90.5)	0.92	6.6 (-111.6, 124.8)	0.87	-52.3 (-108.3, 3.7)	0.07	13.3 (-49.7, 76.2)	0.68
At 41 wks	39.8 (-59.6, 139.2)	0.43	30.2 (-105.7, 166.1)	0.52	-84.7 (-154.1, -15.4)	0.02	45.4 (-31.6, 122.5)	0.25

Summary

The objective of this dissertation is to contribute to the ongoing research on prenatal arsenic exposure and birth outcomes with a view to mechanistically explain these associations by exploring miRNA expression profiles in the placenta. The first part of the dissertation investigates the effects of prenatal arsenic exposure, adolescent marriage, and pregnancy weight gain on preterm birth in Bangladesh. We characterized arsenic exposure in drinking water from 1,181 study participants and in maternal toenail collected 1-month post-partum from 1,142 participants, representing exposure essentially for the whole duration of pregnancy. Because these factors of environmental, societal and maternal health origin often coexist and are highly prevalent in the South and South-East Asia, including Bangladesh, findings from this research may provide insights into effective intervention strategies for reducing preterm birth burden in this region.

In Chapter 2, we critically evaluate the causal association between prenatal arsenic exposure and birthweight in relation to shortening of gestation and intrauterine growth restriction- the two main causal processes of low birthweight. Using quantile causal mediation analysis technique, we were able to estimate the pathway specific effects of arsenic on birthweight across birthweight quantiles. This is the first study to demonstrate that prenatal arsenic exposure was associated with decreased birthweight that vary across the birthweight distribution, and that smaller infants, who are already at higher risk of perinatal mortality and morbidity, are disproportionately affected by arsenic exposure owing to both shortening of gestation and intrauterine growth restriction.

In Chapter 3, we delved into identifying placental epigenetic markers that are associated with birthweight with a view to understand the epigenetic regulation of birthweight in relation to environmental arsenic exposure. In a two-stage study design, we screened 754 human miRNAs in a discovery cohort of 154 participants sampled from both tails of the birthweight distribution using TaqMan qRT-PCR miRNA Array and then conformed expression of 49 top-hit miRNAs in a replication cohort of 364 participants randomly selected from the full study cohort using probe-based qRT-PCR assay. Using causal mediation analysis approach, we identified placenta-derived miRNAs that regulate birthweight not only by determining the length of gestation but also modulating the susceptibility of environmental toxicants on fetal growth. Pathway and network analysis of predicted target genes associated with those miRNAs reveled biological processes including insulin signaling, apoptosis, and regulation of cell proliferation that might play important role in regulating birthweight.

Our research presents a more complete understanding of the associations between prenatal arsenic exposure and birth outcomes by contributing new knowledge to explain the underlying mechanisms of these complex exposure-outcome relations. It is our hope that these findings will be used to increase knowledge and ultimately promote health for vulnerable population that continues to be affected by inorganic arsenic.