Investigating the Association between Arsenic Exposure and Chronic Disease Using Toenail Speciation Biomarkers: A Pilot Study in Atlantic Canada

by

Nathan Kyle Smith

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2017

© Copyright by Nathan Kyle Smith, 2017

Table of Contents

List of Tables ... iv
List of Figures ... v
Abstract ... vi
List of Abbreviations Used ... vii
Acknowledgments ... viii
Chapter 1. Introduction ... 1
1.1 Arsenic as a Global Health Concern ... 1
1.2 Cancer and Arsenic Exposure Risk in Atlantic Canada ... 1
1.3 Arsenic and Type II Diabetes in Canada ... 2
1.4 Toxicity and Metabolism Mechanisms of Arsenic Exposure ... 3
Chapter 2. Research Question and Hypothesis ... 6
Chapter 3. Methods ... 7
3.1 Study Design ... 7
3.2 Atlantic PATH ... 7
3.3 Inclusion/Exclusion Criteria ... 8
3.5 Arsenic Speciation and Covariates ... 9
3.6 Analytical Methods ... 11
3.6.1 Chemicals and Reagents ... 11
3.6.2 Analytical Method development ... 11
3.6.3 Sample Preparation ... 12
3.6.4 HPLC-ICP-MS Analysis ... 13
3.7 Statistical Analysis ... 16
Chapter 4. Results ... 17
4.3 Arsenic Speciation and Heavy Metal Profiles by Disease Group ... 20
4.4 Arsenic Speciation Regression Models for Cancer and Diabetes ... 24
4.5 Regression Models for Heterogeneity in Speciation ... 27
Chapter 5. Discussion ... 30
5.1 Digestion vs Extraction for Optimal Measurement of Metabolic Species ... 30
5.2 Cancer, Diabetes, and Comorbidity Effects ... 30
5.3 Total Arsenic vs Arsenic Speciation ... 31
5.4 Impact ... 31
5.5 Strengths and Limitations ... 32

5.6 Conclusions and Future Research ... 33
Bibliography ... 35
Appendix A: Hypothesis Diagram ... 39
Appendix B: Arsenic Metabolism Diagram ... 40
Appendix C: Sample Size Calculations for Follow-up Study ... 41

List of Tables

Table 1. Operational Parameters of Microwave Digestion Method for Toenails 13
Table 2. HPLC Mobile Phase Gradient 14
Table 3. HPLC Instrument Conditions 15
Table 4. ICP-MS Method Conditions for Arsenic Speciation and Total Metal Analysis ...15
Table 5. Arsenic Species Recovery after Digestion 17
Table 6. Participant Characteristics by Disease Group Group 19
Table 7. Descriptive Characteristics and Frequency Estimates of Participants with and without Missing Data on BMI 20
Table 8. Arsenic Species Proportions by Disease Group 22
Table 9. Average Concentrations of Metals in Toenails by Disease Group 23
Table 10. Diabetes Models for %MMA Exposure Using Exact Logistic Regression 24
Table 11. Crude and Adjusted Combined Cancer Models for %MMA Exposure 26
Table 12. Cancer Subgroup Analyses for %MMA Exposure using Exact Logistic Regression Models 27
Table 13. Regression Models for Influences on %MMA 28
Table 14. Mean Sleep Time and %MMA by Province 29
Table 15. Sample Size Calculations per disease Group for Select Arsenic-related Variables 41

List of Figures

Figure 1. Mean %MMA by Disease Group 21
Figure 2. Toenail Selenium Concentration by Disease Group 24
Figure 3. Arsenic Speciation Hypothesis Diagram 39
Figure 4. Typical Arsenic Metabolism Pathway in Mammals 40

Abstract

Background: Long-term exposure to environmental arsenic has been associated with many chronic diseases including several cancers, diabetes and cardiovascular disease. Urinary studies have implicated arsenic speciation as a risk factor for certain cancers and diabetes. However, no speciation studies to date have used long-term biomarkers of arsenic exposure, or have done direct comparisons between disease groups. The main objective of this study was to investigate the association between arsenic speciation and skin, lung, bladder, and kidney cancer, and type II diabetes, utilizing toenail biomarkers of arsenic exposure.

Methods: A cross-sectional pilot study was carried out using toenail samples and baseline questionnaire data from the Atlantic PATH study. Ten samples were selected from each disease group (skin, lung, bladder, and kidney cancer, type II diabetes, and a healthy reference group), for a total of 60 samples. Arsenic speciation profiles were determined using high performance liquid chromatography (HPLC) and inductively coupled plasma mass spectrometry (ICP-MS). Logistic regression models were used to estimate the associations between arsenic speciation and chronic disease.

Results: No statistically significant differences were found in total arsenic concentration between disease groups. However, statistically significant differences were found with the proportion of monomethylated species (%MMA). For every 1% increase in %MMA, the odds of cancer increased by 114% (OR = 2.14, 95% CI = 1.05, 4.37) when adjusted for BMI, age, and having ever smoked. When analyzed by cancer type, for every 1% increase in %MMA, the odds of skin cancer increased by 68% (OR = 1.68, 95% CI = 1.02, 3.23), the odds of bladder cancer increased by 31% (OR = 1.31, 95% CI = 0.98, 3.08), the odds of lung cancer increased by 46% (OR = 1.46, 95% CI = 1.03, 2.36), and the odds of kidney cancer increased by 259% (OR = 3.59, 95% CI = 1.27 27.12). For every 1% increase in %MMA, the odds of type II diabetes increased by 89% (OR = 1.89, 95% CI = 1.03, 4.78).

Conclusions: The proportion of monomethylated arsenic species (%MMA) was found to be significantly higher in the toenails of individuals with chronic arsenic-related diseases, compared to healthy individuals with similar total arsenic exposure, suggesting that speciation is associated with arsenic-related diseases. These results from indicate that toenail arsenic speciation analysis has the potential to advance our understanding of arsenic pathogenesis and carcinogenicity.

List of Abbreviations Used

AsB	Arsenobetaine
AIC	Akaike Information Criterion
As^{III}	Arsenite
As^{V}	Arsenate
BHCRI	Beatrice Hunter Cancer Research Institute
BMI	Body Mass Index
CI	Confidence Interval
CPTP	Canadian Partnership for Tomorrow Project
CRTP	Cancer Research Training Program
DMA^{III}	Dimethylarsinous acid
DMA^{V}	Dimethylarsinic acid
HPLC	High Performance Liquid Chromatography
iAs	Inorganic Arsenic
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
MAR	Missing at Random
mg	Milligram
mL	Millilitre
mM	Millimolar
MMA^{III}	Monomethylarsonous acid
MMA^{V}	Monomethylarsonic acid
ng	Nanogram
NSE	Nova Scotia Environment
OR	Odds Ratio
PATH	(Atlantic) Partnership for Tomorrow's Health
PMI	Primary Methylation Index
SD	Standard Deviation
SMI	Secondary Methylation Index
µg	Microgram
µL	Microlitre
$(NH_4)_2CO_3$	Ammonium Carbonate
82Se	Selenium (isotope 82)

Acknowledgments

I would like to thank Dr. Jong Sung Kim, Dr. Swarna Weerasinghe, Dr. Kathy MacPherson, and all members of the Health and Environments Research Centre (HERC) Laboratory for their help and support during the development of this project. I would like to give a special thank you to Olivier Collins for his exceptional technical support during the method development process, and also to Crystal Sweeney for her kind and diligent help and support in the lab. Lastly, I would also like to thank Dr. Louise Parker, Dr. Ellen Sweeney and Jason Hicks at Atlantic Partnership for Tomorrow's Health (PATH) for supporting and facilitating this research.

Nathan Smith is a trainee in the Cancer Research Training Program of the Beatrice Hunter Cancer Research Institute, with funds provided by The Terry Fox Research Institute.

Chapter 1. Introduction

1.1 Arsenic as a Global Health Concern

Arsenic exposure has been well documented as a significant risk factor for several types of cancer, including skin (1), liver (2), bladder (3), kidney (4), prostate (5), lung (6) and breast cancer (7), and has been classified as a group 1 known human carcinogen by the International Agency for Research on Cancer (8). Additionally, arsenic is the number one priority substance on the Agency for Toxic Substances and Disease Registry 2015 substance priority list, which ranks substances based on frequency of occurrence, toxicity, and potential for human exposure (9). Arsenic has been consistently ranked as number one on this list since 1997, and before that it was second only to lead (9). Chronic exposure to environmental arsenic has been linked to a number of human diseases affecting multiple organ systems, including several cancers, diabetes, and cardiovascular and neurological disease (10). The greatest concern for such chronic arsenic exposure is through contaminated drinking water. Inorganic arsenic that contaminates water has no color, smell or taste, so testing is needed to identify its presence (11). In fact, arsenic has previously been recognized as the second most important health risk associated with drinking water globally, second only to infectious microorganisms (12). This has been most prominently observed with the high levels of arsenic contamination in drinking water in Bangladesh in 1993, which has been described as the largest mass poisoning in history (13). In contrast with the case of Bangladesh, lower-level arsenic contamination is likely having subtle, long-term effects in other parts of the world.

1.2 Cancer and Arsenic Exposure Risk in Atlantic Canada

Areas with high levels of environmental arsenic are at greater risk for groundwater contamination (11). Some areas of Atlantic Canada have high levels of naturally occurring environmental arsenic (14). Nova Scotia in-particular has high levels of naturally occurring environmental arsenic throughout the province (15). A 2008 soil sample survey found that 56% of sites sampled had arsenic concentrations that exceeded the national soil guidelines of 12 µg total arsenic/g (16). The results of this 2008 survey support previous water quality studies indicating that some areas of Nova Scotia frequently exceed the national water quality guidelines of 10 µg total arsenic/L

(17). Water testing and treatment for contaminants is routinely performed for municipal drinking water supplies to ensure they are within the acceptable health guidelines (17). However, people who live in rural areas are at greater risk because they are more likely to use a private well for their drinking water (18). An estimated 40% of households in Nova Scotia use water from private wells, which are not regulated in the same way as the municipal water supply. A recent study found that 12% of private wells sampled in Nova Scotia had concentrations above the Canadian drinking water safety guideline of 10 µg total arsenic/L (18). The economic cost and burden of responsibility of regular monitoring and testing is left to the homeowner. If homeowners do not follow the provincial monitoring recommendations, they could be at risk of unknowingly consuming contaminated drinking water. Nova Scotia Environment (NSE) recommends testing of well water for arsenic and other chemical contaminants once every one to two years (19). However, compliance is low – a recent survey in Nova Scotia showed that only 12% of private well users surveyed followed provincial recommendations (18).

Geological and environmental factors affecting well water contamination with arsenic have been shown to be the main contributors to the arsenic body burden (amount of arsenic present in tissues: in this case, measured by toenail concentrations) in the Nova Scotian population (20). This is particularly interesting for the current study as Atlantic Canada has the highest age-standardized incidence rates of all cancers combined in the country (21). Atlantic Canada's high cancer rates combined with abundant environmental arsenic and prevalent well water usage make it a prime location for an investigation of arsenic pathogenicity and carcinogenicity.

1.3 Arsenic and Type II Diabetes in Canada

In addition to being associated with several cancers, arsenic exposure and speciation has also been associated with another important and detrimental chronic disease: type II diabetes. The Canadian Diabetes Cost Model has estimated that in 2015, 9.3% of the Canadian population had diabetes, with an additional 22.1% of Canadians with prediabetes (22). Between 2015 and 2025 the prevalence of diabetes is estimated to increase by 44%. This chronic disease has a significant cost associated with it. In 2015 the estimated cost was $14 billion CAD. By 2025 that number is estimated to increase by 25% (to $17.4 billion CAD) (22). Efforts to mitigate this health and financial burden should be focused on prevention. In order to prevent this disease, we need to

understand how it develops, and what factors play an important role. While certain important factors such as obesity have been identified, other less obvious ones such as chronic arsenic exposure may play an important role as well.

Recent studies have shown an association between arsenic speciation in urine and both cancer (1,23) and type II diabetes (24). These studies have shown a different association between arsenic speciation and cancer, compared to arsenic speciation and diabetes. Whereas a higher percentage of monomethylarsonic acid (MMA^{V}) is associated with increased odds of cancer, a higher percentage of dimethylarsenic acid (DMA^{V}) is associated with increased odds of type II diabetes and obesity. Considering these contrasting associations from different studies, a comprehensive study investigating the association of arsenic speciation with both cancer and diabetes simultaneously is needed to help elucidate potential underlying mechanisms of arsenic-related disease etiology.

1.4 Toxicity and Metabolism Mechanisms of Arsenic Exposure

The mechanisms behind arsenic-induced toxicity and carcinogenicity have not been fully elucidated, but there is evidence to suggest that metabolism is an important factor. The toxicity and bioavailability of arsenic is highly dependent on its form or species. Different forms of arsenic can induce different effects on the human body, ranging from relatively harmless to very toxic (11,25). The human body metabolizes arsenic by successive methylation in the liver, and the metabolites are ultimately excreted in urine (which serves as a short-term indicator) or deposited in keratin rich tissues such as hair or nails (which serve as longer-term indicators) (26). However, metabolism of arsenic is incomplete and a mixture of inorganic and mono and dimethylated species are deposited/excreted (27). Previously, it was generally agreed upon that methylation was a detoxification process; however, more recent findings suggest that methylation could activate toxic and carcinogenic effects of arsenic (25).

Individuals may differ in their ability to process and metabolize arsenic, and previous studies have found associations between the extent of metabolism and different arsenic-related diseases including skin cancer, bladder cancer, and diabetes (1,23,24). A higher percentage of monomethylated species (indicating incomplete metabolism) in urine has been associated with skin and bladder cancer (1,23). Conversely, a lower percentage of monomethylated species has been associated with

diabetes and obesity (24). These and other similar findings indicate that arsenic metabolism plays a role in pathogenesis. However, further epidemiological evidence is needed to confirm these associations. Additionally, these associations have not been explored in other important arsenic-related diseases such as lung cancer, which has also been linked to drinking water exposure (25).

Most studies assessing speciation have focused on analysis of urine, in which arsenic has a short residence time (3 - 4 days), and which is thus better suited for assessing short-term, acute exposure (28). Since many arsenic-related diseases arise from chronic exposure, and can develop over the course of several years, measuring longer-term biomarkers may prove to be a more relevant assessment. Toenails have been shown to be more reliable as markers of long-term exposure than hair or fingernails, as they are less prone to contamination and washout effects from things such as shampoo, nail polish, or regular handling of items (29). Additionally, toenails grow more slowly than hair or finger nails, so they can provide a longer period (approximately 2 - 12 months) of exposure data (28). While toenails have been demonstrated to be useful indicators of long-term arsenic exposure, arsenic speciation analysis of different disease groups has not yet been conducted, and could provide valuable insights into how arsenic causes disease.

Toenails are a relatively new biomarker for environmental heavy metal exposure. They have several advantages: they are non-invasive to collect, provide a longer-term exposure estimate than conventional biological samples (2 - 12 months, compared to a few days from urine), and the samples are very stable and can be stored at room temperature for long periods of time (28). These sample properties also make biobanking for future studies a very feasible option for researchers and clinicians. While a handful of studies have shown that toenails can provide a good measurement of long-term exposure via contaminated drinking water, few studies have looked at arsenic speciation in nail samples, and none have looked at toenail arsenic speciation in the context of cancer or diabetes. The few studies that have investigated specific cancers or disease groups have only used short-term urinary measures (1,3,23,24,30). The arsenic speciation studies that have used the more appropriate toenail biomarker, have only compared these findings to arsenic species in hair, nails, urine, water and food - but not in the context of disease (26,31,32). To the best of our knowledge, this is the first study investigating the associations between arsenic exposure, various cancers, and type II

diabetes using toenail arsenic speciation analysis, a potentially important, more relevant biomarker for chronic diseases.

Chapter 2. Research Question and Hypothesis

In Atlantic Canadians aged 35 - 69 who use private well water, are certain characteristic arsenic speciation profiles of toenail biomarkers associated with increased odds of having skin, lung, bladder, or kidney cancer, or type II diabetes mellitus?

To address this question, we hypothesize that arsenic speciation may mediate the effect of arsenic exposure on the development of arsenic-related diseases, thus individuals with different arsenic-related diseases will have differential toenail arsenic speciation profiles. Specifically, we hypothesize that there is a higher proportion of monomethylated arsenic (MMA) species in individuals with cancer and a higher proportion of dimethylated arsenic (DMA) species in individuals with diabetes mellitus, when compared to healthy controls without either disease (refer to Figure 3, Appendix A).

A limitation to testing this hypothesis is that we do not know what species of arsenic individuals are ingesting. Thus, we will not know whether individual differences are a result of differential metabolism, or differential exposure. This hypothesis was tested assuming that the majority of exposure is in the form of inorganic arsenic via drinking water, as previous work has demonstrated that drinking water is the main contributor to arsenic body burden in the Nova Scotian population (20). Given this assumption, results should indicate metabolic differences rather than exposure differences.

The main goal of the study is to advance arsenic-speciation research in the context of chronic disease. To achieve this goal, we have established the following objectives:

1. Characterize toenail arsenic speciation profiles of individuals by disease category: skin cancer, lung cancer, bladder cancer, kidney cancer, type II diabetes mellitus, and a healthy reference group.
2. Determine if variation in arsenic speciation is significantly associated with cancer and/or diabetes.
3. Identify factors associated with heterogeneity in speciation.

Chapter 3. Methods

3.1 Study Design

This project used a cross-sectional study design with arsenic speciation as the primary exposure of interest (measured via toenail samples, indicating exposure over the year or so prior to collection), among those with and without cancer and/or diabetes. We analyzed exposure data and biological samples collected from the Atlantic Partnership for Tomorrow's Health (Atlantic PATH) study. Toenail samples from the selected groups were analyzed using High Performance Liquid Chromatography (HPLC) coupled with Inductively Coupled Plasma-Mass Spectrometry (ICP-MS) to determine concentrations of each arsenic species and other toxic metal as co-contaminants that have previously been found in nails, such as lead, cadmium, chromium, and nickel (28). Additionally, we measured levels of other heavy metals such as copper, zinc, iron, and aluminum. Selenium was also measured, as previous studies have found that selenium can negate the negative effects of arsenic exposure (33). For arsenic speciation analysis, three forms were measured to assess exposure and estimate the extent of metabolism: 1) inorganic arsenic (arsenite, As^{III} plus arsenate, As^{V}), 2) monomethylarsonic acid (MMA^{V}), and 3) dimethylarsinic acid (DMA^{V}).

3.2 Atlantic PATH

Atlantic PATH is part of the Canadian Partnership for Tomorrow Project (CPTP), a large a multi-centred prospective cohort study spanning eight Canadian provinces. The CPTP is comprised of five cohorts including Atlantic PATH, Alberta's Tomorrow Project, the Ontario Health Study, the BC Generations Project, and Quebec's CARTaGENE. CPTP has recruited more than 300,000 participants from across the country to examine the biological and environmental factors that lead to the development of cancer and other significant chronic diseases. Participants will be followed for 30 years. The only exclusion criteria for participation was age (participants must be between 35 - 69). Data collection took place between 2009 and 2015. All participants provided informed consent as part of their participation in the study. Recruitment was done through advertising, media coverage, community and workplace events, incentive programs such as Airmiles, designated community "champions" who encouraged their friends and families to participate, and invitations from the provincial health insurance provider (Nova Scotia

only). Participants provided data and samples in person at assessment centres, and had the option to complete an online or paper questionnaire which could be mailed or delivered in person at an Atlantic PATH office.

In total, Atlantic PATH has collected data from 31,173 individuals aged 35-69 years across the four Atlantic provinces (Nova Scotia, New Brunswick, Prince Edward Island, and Newfoundland and Labrador). Participant data includes demographic information, personal and family health history, lifestyle factors, body measurements, as well as drinking water samples, and biological samples including blood, urine, and toenails. Questionnaire data and biological samples (toenails only) collected at baseline were used for the present study.

3.3 Inclusion/Exclusion Criteria

Participants included in the disease groups must have a current diagnosis (at time of baseline data collection) of type II diabetes mellitus, or have been previously diagnosed (at any age) with primary skin, lung, bladder, or kidney cancer. Participants in the healthy reference group must have never been diagnosed with type II diabetes mellitus, or any type of cancer. All participants must *i) live in Atlantic Canada, ii) receive water from a private source (i.e., not municipal treated water), and iii) have not moved residence within the past ten years*. The healthy comparison group was randomly selected from a group of 8000 Nova Scotians that had previously been cleared as "healthy" from another Atlantic PATH study. For the disease group selection, a few exceptions were made to get 10 samples per group, given the small numbers of participants with certain cancers who provided toenail samples. All participants in the disease groups were first randomly selected from the entire Atlantic PATH cohort (all provinces) based on the above criteria (if enough were available). If there were not enough participants available to meet the target sample size of 10, additional participants were randomly selected by ignoring the limiting exclusion criterion. After selection of the 60 samples, four exceptions were made for a total of three individuals. One individual used municipal water, the second had a residence dwell time of less than ten years, and the third used municipal water *and* had a dwell time of less than ten years. All data used for the inclusion and exclusion criteria, including health history, were obtained via the Atlantic PATH questionnaire.

These criteria allow for a potentially long exposure window in which consistent exposure necessary for disease to develop is conceivable. Participants are likely to have similar and consistent exposure to private unregulated well water, maximizing the chance of obtaining arsenic-exposed individuals. We wished to obtain study participants who have been exposed to similar levels of arsenic, so we can test if the speciation that constitutes the total arsenic is associated with disease. If we were to measure unexposed or lesser-exposed individuals, it is likely that the levels of individual species would have been below the limit of detection, yielding only total arsenic concentration (if any) from those individuals.

3.5 Arsenic Speciation and Covariates

Arsenic can occur as trivalent or pentavalent species, denoted by the superscript III or V. Trivalent species are considered more toxic. Methylated trivalent species (MMA^{III}, DMA^{III}) are also less stable, and act as intermediates in the metabolic pathway (27,33). Thus, in speciation analysis generally both types of inorganic species (As^{V}, As^{III}) are measured, but only the pentavalent methylated species are measured (MMA^{V}, DMA^{V}). Arsenobetaine is also commonly measured in speciation studies. This organic form of arsenic found in seafood is relatively non-toxic and is rapidly excreted within a few days (11). As a result of its non-toxic nature it is largely clinically insignificant, and must be differentiated from the remaining arsenic. It is usually measured and taken into account when analyzing urine so the total amount of clinically significant arsenic is not overestimated (11,31). However, as a result of its rapid clearance from the body, it has not been detected in nails and hair like the other forms of arsenic. In this study, we used arsenobetaine as an internal standard to ensure accurate calibration and quantification during HPLC/ICP-MS analysis.

The classical metabolic pathway is as follows: $As^{V} \rightarrow As^{III} \rightarrow MMA^{V} \rightarrow MMA^{III} \rightarrow DMA^{V} \rightarrow DMA^{III}$ (27,34). A detailed diagram of arsenic metabolism is shown in Figure 4, Appendix B (modified from Aphosian et al. 2004 (34)). Individual differences in methylation capacity lead to incomplete transformation, and a mixture of inorganic and mono and dimethylated species are excreted (27). Previous studies using urinary analysis have used the primary methylation index (PMI) and secondary methylation index (SMI) as surrogate measures of methylation capacity (1,23,27). Simplified, arsenic metabolism can be thought of as a two-step process: 1) the conversion of inorganic

arsenic (iAs) to MMA, and 2) the conversion of MMA to DMA (iAs→MMA→DMA). The PMI is defined as the ratio of MMA^{V} to inorganic arsenic (As^{III} + As^{V}) (step 1), and the SMI is defined as the ratio of DMA^{V} to MMA^{V} (step 2). These measures give an estimate of the extent of methylation and metabolism (23). However, these measures may not be directly relatable to toenail concentrations, as the proportions and dominant species between urine and nails are quite different (1,35). The differences between arsenic species in urine versus toenails is likely due to the differences in the various species' affinity for excretion or storage. However, studies have shown that urine and toenail species are significantly positively correlated (36). Therefore, while the measures of PMI and SMI may not be the same for toenail samples, it is worth exploring to see if the same associations are observed, or if others arise.

The exposure in this study is not the amount of arsenic, but rather the composition of the arsenic. The exposure measures are the proportions of each form of arsenic contributing to the total arsenic concentration in toenail samples: iAs (As^{V} + As^{III}), MMA^{V}, DMA^{V}. From these measures, the PMI and SMI were calculated and used as exposures in the regression model. In this study, only the pentavalent organic forms were measured. For simplicity, the shorthand MMA and DMA used in this the remainder of the document are referring to the pentavalent species (MMA^{V}, DMA^{V}).

The Atlantic PATH questionnaire has collected information on several potentially significant variables which were included in the analysis. These include body mass index (BMI) (current), sex, age, province, length of stay at current residence, tobacco use (amount, frequency, duration, age started), environmental tobacco smoke (frequency, duration), family history of cancer and diabetes (parents and siblings), use of tanning beds (in the past year), alcohol use (amount, frequency and type in the past year), physical activity (frequency and intensity in the last seven days), age of cancer diagnosis, age of diabetes diagnosis, and chronic co-morbidities. All of these measures are self-reported, and thus are subject to recall bias. Additionally, some measures are current (e.g., BMI), in the past week (physical activity), or in the past year (alcohol). A potential limitation of using these measures is that they may not be representative of previous (or "usual") exposure levels. Other metal co-contaminants in toenail samples will also be measured during the arsenic speciation analysis. Lead, copper, cadmium and selenium concentrations were tested in the models, as previous literature has noted interactions with arsenic (33). All the variables tested were potentially associated with

either the exposure (e.g., metabolic differences as a result of tobacco use, age, sex), or the outcome (e.g., increased risk of disease from family history, tobacco use, BMI) or both.

3.6 Analytical Methods

3.6.1 Chemicals and Reagents

All analytical chemicals including standards and mobile phase used in this study were trace metal grade. Arsenic III and V standards (1,000 mg/L), dimethylarsinic acid (cacodylic Acid, 500 mg), arsenobetaine (50 mg), ammonium carbonate (50 g) and hydrogen peroxide were supplied from Sigma-Aldrich (MO, USA). Disodium methyl arsonate hexahydrate (MMA standard, 500 mg) was purchased from Chem Service (PA, USA). A multi-element standard was purchased from Inorganic Ventures (IV-ICPMS-71A, VA, USA). Nitric acid was supplied from Fisher Scientific (NH, USA). All aqueous solutions were prepared using high purity, deionized water (Milli-Q water system, Millipore, Fisher Scientific, NH, USA). All analytical standards were matrix-matched to the samples (1% HNO_3, 5% H_2O_2). Note, all percentage HNO3 concentrations listed in the methods are referring to percentage of concentrated stock solution (70%), e.g., 1% of the 70% concentrated HNO3 stock solution (approximately 157 mM).

3.6.2 Analytical Method development

Analytical methods for arsenic speciation analysis using the HPLC/ICP-MS in HERC Laboratory needed to be established and validated prior to the speciation analysis of toenail samples. Previous studies using toenails have shown poor extraction efficiency for arsenic speciation. Mandal et al. (2003) reported an average extraction efficiency of 62.7% (37), while Button et al. (2009) reported and average of 53% (35). This is largely a result of the fact that As^{III} is easily oxidized to As^{V}, thus to accurately quantify As^{III} and As^{V} extraction procedures cannot have oxidizing conditions. An alternative to this is to measure summary inorganic arsenic (As^{III} + As^{V}, all measured as As^{V} after As^{III} is oxidized). In this case, the analysis cannot differentiate the two forms of inorganic arsenic, however it allows more options in the sample preparation method, and potentially better recovery for the measured analytes. In this study, we tested the feasibility of acid digestion for arsenic speciation analysis.

A standard wet digestion method for toenails uses concentrated nitric acid, which requires too much dilution for HPLC/ICP-MS speciation analysis (e.g., final concentration of nitric acid in samples needs to be lower than 1%) leaving the concentrations of arsenic species below the detection limit. Thus, we needed to establish a digestion method to use a lower concentration of nitric acid (10% vs the usual 50%). To test if this dilute acid digestion performed as well as typical acid digestions (in terms of releasing bound arsenic inside toenail into solution), two equal toenail samples were digested and compared for total arsenic concentration; one with the standard acid concentration (50% H_2O_2 and 50% concentration HNO_3), and one with the modified mixture proposed for this study (50% H_2O_2 and 10% HNO_3).

After verifying the two digestion methods perform similarly, the next step was to test if certain arsenic species were being lost or modified during digestion. To ensure species stability through digestion, analyte standards were digested and compared to non-digested equivalents to determine recovery: Enough of each arsenic species (AsB, As^{III}, As^{V}, MMA, DMA) to get a final concentration of 1 µg/L after dilution were spiked into a reagent blank (50% H_2O_2 and 10% HNO_3) and digested. A matching non-digested spiked blank was used to determine the recovery after digestion (% Recovery = [digested spike/blank spike]). We used the modified digestion method (10% nitric acid) for toenails in this study because it provides good recovery of all target arsenic species (refer to Table 5).

3.6.3 Sample Preparation

Toenail samples were washed prior to analysis using a modified protocol previously described by Slotnick et al. (38). Visible contamination was manually removed (scraped with a quartz fragment). The samples were then weighed prior to washing and drying to obtain the wet mass. Samples were weighed with a target of approximately 55 mg, to ultimately get a dry mass of approximately 50 mg for digestion. This amount is based on the preliminary data and allows for accurate speciation analysis. However not all samples had sufficient mass. In these cases, the entire sample was used. Samples were then sonicated in acetone for five minutes, consecutively rinsed with acetone and Milli-Q water, sonicated in Milli-Q water for 10 minutes, and rinsed with Milli-Q water three times before being dried in an oven for twelve hours at 105°C. Dried samples were

then weighed again to correct for moisture content (calculation for percentage moisture content: (wet mass / dry mass - 1) x 100).

A microwave-assisted digestion was used to digest the toenails samples using a Discovery SPD Microwave Digester (CEM Corporation, NC, USA). Briefly, toenails were added into 10 mL quartz pressure vials (CEM Corporation, NC, USA) with 1 mL solution containing 50% hydrogen peroxide and 10% nitric acid, and a disposable magnetic PTFE coated stir-bar (10 x 3 mm, VWR, PA, USA). A digestion reagent blank (toenail-free reagent solution) was run with every batch to check for trace contamination during sample preparation and digestion. A water clean sample was run after every 5 samples to keep the system clean during the run. Following digestion, the samples were allowed to cool in an autosampler rack for at least 10 minutes. Digested samples were diluted ten-fold with Milli-Q water to obtain approximately 1% nitric acid concentration (1% of concentrated HNO_3 solution - approximately 157 mM), and syringe filtered through a 0.45 µm cellulose acetate filter before analysis. Operational parameters of the microwave-assisted digestion method for toenails are shown in Table 1.

Table 1. Operational Parameters of Microwave Digestion Method for Toenails

Parameter	Value
Control Style	Ramp to Temperature
Pressure Mode	Organic
Temperature	165 °C
Ramp Time	4:00 min
Hold Time	4:00 min
Pressure	400 PSI
Power	300 W
Stirring	Medium
Sample Weight	50 mg
Reagents	100 µL HNO_3, 400 µL H_2O, 500 µL H_2O_2
Total reagent volume	1 mL

3.6.4 HPLC-ICP-MS Analysis

Arsenic species were separated using an inert/biocompatible HPLC (Spectra System, Thermo Fisher Scientific, MA, USA). An anion exchange column (Dionex AS7, Thermo

Fisher Scientific, MA, USA) was used with an ammonium carbonate (99.999% trace metal basis, Sigma, MO, USA) mobile phase gradient ranging from 20 - 200 mM (Table 2) to separate the arsenic species. For quantification, the outlet of the column was connected directly to the inlet of the ICP-MS with PEEK tubing (ICAP-Q, Thermo Fisher Scientific, MA, USA). Full instrument conditions for both HPLC and ICP-MS are listed in Tables 3 and 4, respectively. A calibration curve was constructed using a mixed standard solution of the three arsenic species using the following concentrations: 0.02 µg/L, 0.05 µg/L, 0.1 µg/L, 0.5 µg/L, 1.0 µg/L, and 2.0 µg/L. This calibration curve enabled us to determine unknown concentrations in the biological samples, within the trace-metal range described by previous literature (31,35). Internal standardization was also used to correct for instrumental drift. Using an internal standard (IS) compound similar to, but not identical to our target analytes ensures that factors which affect the response of one should similarly affect the response of the other. Internal standardization uses the ratio of response between the analyte and the IS to correct for drift. Arsenobetaine was used as the internal standard, and was added to all samples and calibration standards at a concentration of 0.05 µg/L. Arsenobetaine was chosen as the internal standard because it is similar to, but not identical to the other arsenic compounds, and is not found in toenails. A continuing calibration verification quality control check standard was analyzed every 10 samples to ensure the quality of data.

Table 2. HPLC Mobile Phase Gradient

Time (min)	H2O (%)	20 mM $(NH_4)_2CO_3$ (%)	200 mM $(NH_4)_2CO_3$ (%)
0	75	25	0
0.5	45	35	20
0.6	0	0	100
2.1	0	0	100
2.2	75	25	0
6.2	75	25	0

Method modified from Ellingson et al. (39) to allow for faster analysis while not needing to separate out As^{III}.

Table 3. HPLC Instrument Conditions

Injection Volume	25 µL
Flow Rate	0.4 mL/min
Flush Volume	1000 µL
Sample Loop Size	100 µL
Injection Type	Push Loop
Analytical Column	Dionex AS7, 2 x 250 mm
Guard Column	Dionex AG7, 2 x 50 mm

Table 4. ICP-MS Method Conditions for Arsenic Speciation and Total Metal Analysis

	Arsenic Speciation	**All Metals**
Gas	Argon (99.999%)	
Mode	Kinetic energy discrimination (KED)	
Collision Cell Gas	Helium (99.999%)	
Instrument Setup	HPLC	Autosampler
Dwell Time	0.5 seconds	0.1 seconds*
Internal Standard	Arsenobetaine	Scandium
Main Runs	N/A	3
Quality Control Frequency	Every 10 Samples	

*All elements were analyzed with 0.1s dwell time except arsenic, which was analyzed with 0.5s to match the speciation analysis.

For determination of total arsenic and other metals, a separate ICP-MS analysis was performed. A calibration curve was generated for 18 metals (Ag, Al, As, Cd, Ce, Co, Cr, Cu, Fe, Mn, Ni, Pb, Rb, Se, Sr, U, V, Zn) with the following concentrations: 0.05 µg/L, 0.1 µg/L, 0.5 µg/L, 1.0 µg/L, and 5.0 µg/L. All elements were analyzed in KED mode with 0.1s dwell time (with the exception of arsenic, which was kept at 0.5s), and scandium was used as the internal standard (Table 4). A quality control check was run every 10 samples.

All final toenail metal concentrations including arsenic species were calculated using the toenail sample dry mass (measured after washing samples and drying at

105°C for 12 hours) to adjust for moisture content. All concentrations were reported as nanograms of arsenic per gram of toenail (ng/g).

3.7 Statistical Analysis

A variety of statistical techniques were employed to achieve our objectives. A one-way analysis of variance (ANOVA) of MMA percentages was used to determine if the disease sub-groups have significantly different speciation patterns, and a Dunnett's pairwise multiple comparison of means was used to compare mean MMA percentages of each cancer group to that of the healthy reference group (Objective 1). Multiple logistic regression models using disease/no disease as the outcome, and arsenic speciation (%MMA) as the exposure of interest were used to determine if arsenic speciation was significantly associated with increased odds of having cancer or diabetes, while adjusting for significant confounders: An "all cancers combined" analysis was performed, in addition to subgroup analyses by cancer type. The cancer subgroup and diabetes analyses were performed using exact logistic regression models, due to small numbers of participants with those specific conditions ($N < 10$), (Objective 2). Exact logistic regression was chosen because it is a method designed to work with small samples of data to achieve high accuracy (40). Biologically plausible interaction terms were also tested for significance in the multiple logistic regression models. To identify factors associated with heterogeneity in speciation, we used a multiple regression model using the percentage of MMA as the outcome, and dataset covariates (disease status, demographic information and other exposures) as the predictor variables (Objective 3).

The strategy for fitting best regression models for objectives 2 and 3 was as follows: 1) Backward stepwise regression starting with a full model, removing everything with $p > 0.1$, one variable at a time. 2) The significant variables were kept, and a few of the variables with p-values closest to the 0.1 cut off that were dropped, were sequentially added back into the model. 3) Sequential models were compared, and the goodness of fit of each model was determined by comparing Akaike information criterion (AIC). The final model included variables that resulted in the lowest AIC score. All statistical analyses were performed with Stata/SE 13.

Chapter 4. Results

4.1 Method Development and Laboratory Analysis

Comparing the standard acid digestion (50% H_2O_2 and 50% concentration HNO_3) with the modified mixture proposed for this study (50% H_2O_2 and 10% HNO_3) resulted in mass balance of 95.5% (SD=4.9, n=2) between the two methods, indicating that the two performed similarly, and that the modified method was suitable for this study. Spike recovery tests with the modified digestion method showed that all arsenic species remained stable through the digestion process, with the exception of As^{III}, which was fully oxidized and almost completely recovered as As^{V} (Table 5).

Table 5. Arsenic Species Recovery after Digestion

	Spike Recovery	**SD (n=2)**
MMA	97.9%	4.1%
DMA	99.3%	9.2%
As^{III}*	98.4 %	0.7%
As^{V}	110.3% **	N/A
iAs (As^{III} + As^{V})*	108.1%	0.8%

1 µg/L standards were spiked and digested. A matching non-digested spiked blank was used to determine % recovery after digestion. Calculation: % Recovery= (digested spike/blank spike).
*Recovered as As^{V} (As^{III} completely oxidized to As^{V} under digestion conditions).
**n=1.

After validation of the analytical method, 60 toenail samples from Atlantic PATH (10 from each disease category: skin, lung, bladder, and kidney cancer, type II diabetes and healthy controls) were analyzed for a panel of 18 metals, with 46 of the samples allowing for quantifiable arsenic speciation analysis. The average moisture content of the 60 toenail samples was 11.30% (SD = 3.00%). A few of the samples provided were less than 15 mg (4 samples), and thus were not suitable for speciation analysis. Samples with recoveries off by $\geq$30% were dropped (9 samples). Additionally, one sample had total arsenic content too low to determine speciation (15.19 ng/g). Interestingly, three of these samples with poor recoveries had similar characteristic chromatograms. The reason for this is unknown, but is likely due to some sort of interference by something released from the toenails after digestion. The column had to be rinsed by injecting several blanks after each of these samples to return the baseline to normal. The average

extraction efficiency (measured by mass balance of combined As species/total As) of the included 46 speciation samples was 97.6 % (SD = 14.0).

4.2 Participant Characteristics

The mean age of the sample was 59.25 years, and 56.67% were women (Table 6). All healthy individuals were from Nova Scotia as a result of sampling from a different participant pool for the healthy group and disease groups. All lung cancer cases got their drinking water from a dug private well, and skin cancer cases had the longest average residence dwell time. There were no other significant differences between groups. Additionally, BMI had the most missing observations of all variables tested (28.33%). All sample population characteristics are summarized in Table 6.

Since BMI had so many missing observations, we tested if these observations were missing in a systematic fashion by comparing the characteristics of those missing BMI, to those with a BMI observation (Table 7). No significant differences were found between the two groups, so we moved forward under the assumption that the data are missing at random (MAR). Hotdeck imputation was performed for BMI so we did not have to drop those data points from the regression models. Imputation was based on the two variables most highly correlated with BMI that had no missing values, i.e., cancer history and age.

Table 6. Participant Characteristics by Disease Group

Variable		N (Missing)	Healthy % (Mean)	Skin Cancer % (Mean)	Bladder Cancer % (Mean)	Lung Cancer % (Mean)	Kidney Cancer % (Mean)	Type II Diabetes % (Mean)	P value	Total % (Mean)
Age		60 (0)	(55.9)	(64.3)	(60.8)	(57.8)	(59.1)	(59.2)	0.1549	(59.52)
Sex	Male	60 (0)	30	30	60	40	30	70	0.283	43.33
	Female		70	70	40	60	70	30		56.67
Province	NS	59 (1)	100	11.11	20	50	0	20	<0.001*	33.90
	NB		0	77.78	60	30	40	20		37.29
	NL		0	11.11	10	0	50	40		18.64
	PEI		0	0	10	20	10	20		10.17
Ever Smoked	No	59 (1)	60	55.56	40	10	30	37.29	0.209	37.29
	Yes		40	44.44	60	90	70	62.71		62.71
BMI		43 (17)**	(36.05)	(29.64)	(29.29)	(27.24)	(29.13)	(34.30)	0.0841	(31.20)
Sleep Time (min)		59 (1)	(429)	(474)	(546.11)	(408)	(450)	(429)	0.5264	(440.76)
Water Source	Well-Dug	60 (0)	70	30	70	100	50	70	0.019	65.00
	Well-Drilled		20	30	10	0	40	20		20.00
	Surface		10	40	0	0	10	10		11.67
	Municipal		0	0	20	0	0	0		3.33
Residence Time (years)		60(0)	(24.5)	(42.3)	(25.6)	(26.5)	(31.6)	(45)	0.0004	(32.58)

Cancer groups including comorbid diabetes (n = 4, 2 in bladder 1 in skin, 1 in kidney), n = 10 per disease group. Categorical variables reported as % frequency, P-value indicates Pearson chi-squared significance. Continuous variables reported as means, P-value indicates one way ANOVA significance.

*All healthy individuals from Nova Scotia as a result of sampling from a different participant pool for the healthy group.

**Large proportion of participants with missing values for BMI.

Table 7. Descriptive Characteristics and Frequency Estimates of Participants with and without Missing Data on BMI

Variable		Not Missing % (Mean)	Missing % (Mean)	P-value
Age		(60.09)	(58.06)	0.3261
%MMA		(5.33)	(6.52)	0.2694
Sex	Male	80.77	19.23	0.171
	Female	64.71	35.29	
Province	NS	65.00	35.00	0.284
	NB	72.73	27.27	
	NL	90.91	9.09	
	PEI	50.00	50.00	
Status	Healthy	70.00	30.00	0.115
	Skin Cancer	100.00	0.00	
	Bladder Cancer	44.44	55.56	
	Lung Cancer	60.00	40.00	
	Kidney Cancer	66.67	33.33	
	Type II Diabetes	90.00	10.00	

No significant differences, data are assumed missing at random (MAR). P-values are results from Student's t-test (age and %MMA) and one-way analysis of variance (sex, province, status).

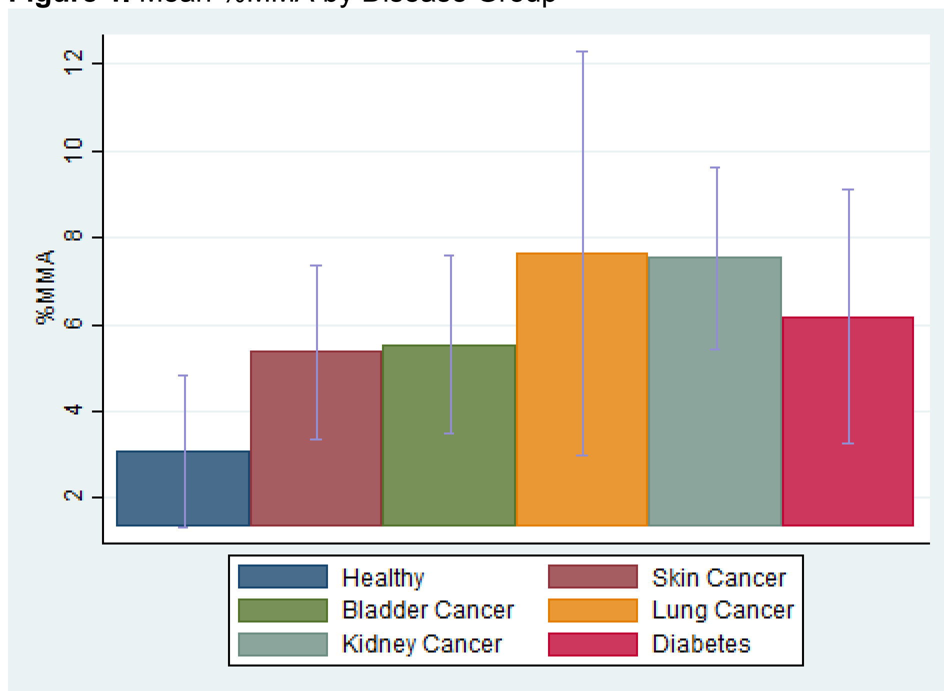
4.3 Arsenic Speciation and Heavy Metal Profiles by Disease Group

A one-way analysis of variance (ANOVA) was performed for all speciation variables using all six groups, and was only significant for %MMA ($p = 0.0430$, Table 8), and total MMA ($p = 0.0002$, Table 9). An ANOVA comparing %MMA of the four cancer groups was not significant ($p = 0.2436$), thus we are unable to reject the null hypothesis that they are all the same. This suggests that we may pool the four groups for an "all cancers" analysis. A Dunnett's pairwise comparison for %MMA was significant for lung cancer ($p = 0.025$) and kidney cancer ($p = 0.023$), indicating that the mean %MMA for lung and kidney cancer are significantly different (higher) than that of the healthy group. A combined analysis of all cancers vs the healthy group using a two-sample t-test indicated that the mean %MMA was significantly different (higher) in the combined cancer group, compared to the healthy group ($p = 0.0084$).

For total metal concentrations, Dunnett's test indicated that average selenium (82Se) in the lung cancer group was significantly different (lower) than the healthy group ($p = 0.016$, Table 9). Total MMA and DMA were also significantly different (higher) in the lung cancer group, compared to the healthy group ($p < 0.001$, $p = 0.031$, respectively, Table 9). No significant differences in means were found for the other metals measured, including total arsenic. Interestingly, lung cancer was the only group with a higher total arsenic concentration than the healthy group, although the difference was not statistically significant (Table 9).

Aside from the statistically significant results, Tables 8 and 9 also show some interesting trends. While not statistically significant, possibly as a result of the small sample sizes, trends in Table 8 indicate that on average, healthy individuals had the lowest %DMA, the highest %iAs, the lowest PMI, and the highest SMI, agreeing with previous literature (Table 8). Additionally, on average healthy individuals had the highest selenium concentration (82Se, Table 9), and individuals with diabetes had the highest %DMA (Table 8).

Figure 1. Mean %MMA by Disease Group



N = 9, 7, 5, 7, 8, 8 from left to right respectively. Means not including comorbid type II diabetes. Error bars indicate 95% confidence intervals.

Table 8. Arsenic Species Proportions by Disease Group

	Healthy	Skin Cancer	Bladder Cancer	Lung Cancer	Kidney Cancer	Type II Diabetes	P Value (ANOVA)	Total Sample	All Cancers Combined
%MMA	3.09 (1.32, 4.85) N = 9	5.53 (3.82, 7.23) N = 8	4.62 (1.78, 7.46) N = 6	**7.66*** (3.00, 12.32) N = 7	**7.54*** (5.44, 9.64) N = 8	6.18 (3.24, 9.11) N = 8	0.0430	5.72 (4.71, 6.72) N = 46	**6.41**** (5.15, 7.67) N = 29
%DMA	9.94 (4.77, 15.12) N = 9	13.84 (8.02, 19.66) N = 8	10.65 (2.31, 18.99) N = 6	14.56 (2.71, 26.41) N = 7	12.63 (9.08, 16.19) N = 8	14.68 (8.19, 21.17) N = 8	0.7741	12.71 (10.40, 15.01) N = 46	13.02 (9.95, 16.09) N = 29
%iAs	86.97 (80.59, 93.94) N = 9	80.63 (73.78, 87.50) N = 8	84.73 (74.89, 94.57) N=6	77.78 (61.40, 94.17) N = 7	79.82 (74.72, 84.93) N = 8	79.15 (70.01, 88.28) N = 8	0.4890	81.58 (78.46, 84.69) N = 46	80.57 (76.49, 84.65) N = 29
PMI	0.048 (0.026, 0.070) N = 7	0.071 (0.046, 0.095) N = 8	0.069 (0.031, 0.107) N = 5	0.121 (0.013, 0.229) N = 7	0.097 (0.064, 0.130) N = 8	0.086 (0.027, 0.145) N = 8	0.3437	0.083 (0.061, 0.105) N = 43	0.090 (0.065, 0.116) N = 28
SMI	3.06 (1.06, 5.05) N = 7	2.73 (1.65, 3.81) N = 8	2.37 (1.02, 3.72) N = 4	1.75 (1.27, 2.24) N = 7	1.75 (1.22, 2.29) N = 8	2.41 (1.73, 3.10) N = 8	0.2445	2.34 (1.96, 2.72) N = 42	2.13 (1.76, 2.51) N = 27

Cancer groups including comorbid diabetes (n=2; 1 skin cancer and 1 bladder cancer).
*Significantly different than healthy (α=0.05) using Dunnett's pairwise comparison: allows for significance correction after multiple comparisons to the healthy reference group (reduced level of significance = α/number of comparisons). Note: 95% confidence intervals are only for the individual mean per group, while Dunnett's test is for the difference between the means.
**Significantly different compared to the healthy reference group using students t-test (α=0.05). p= 0.0084.

Table 9. Average Concentrations of Metals in Toenails by Disease Group

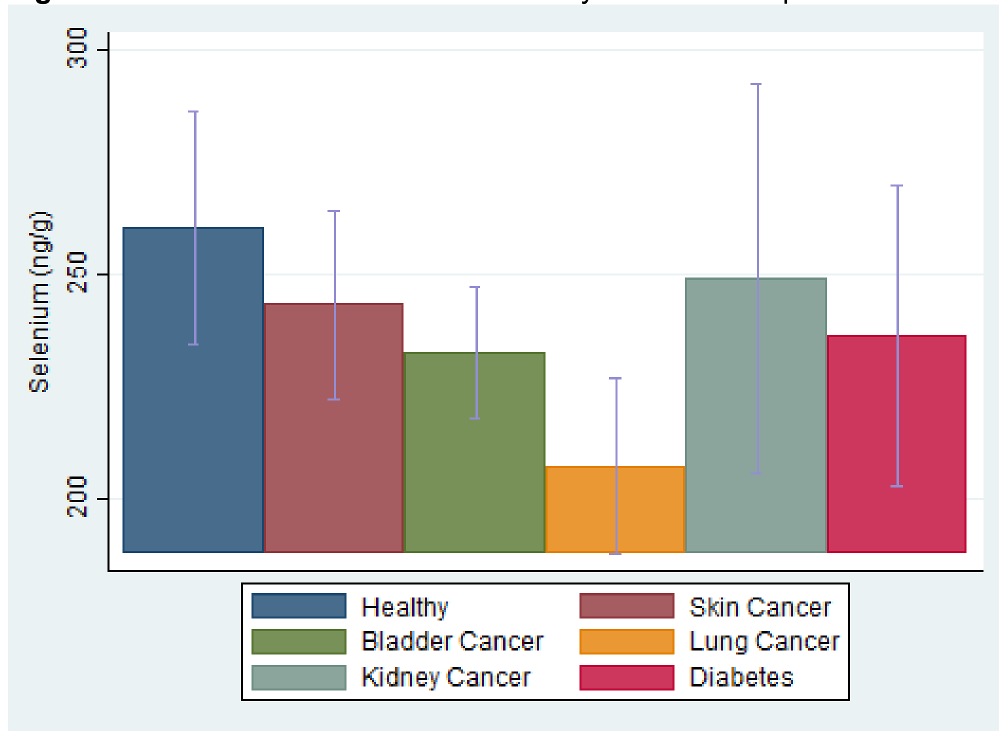
	Healthy	Skin Cancer	Bladder Cancer	Lung Cancer	Kidney Cancer	Type II Diabetes	P Value (ANOVA)	Total Sample	All Cancers Combined
82Se (ng/g)	260.39 (234.63, 286.16) N = 10	243.29 (222.49, 263.97) N = 10	232.53 (217.84, 247.22) N = 10	**207.50*** (187.99, 227.0) N = 10	249.22 (205.92, 292.53) N = 10	236.28 (202.88, 269.68) N = 10	0.0766	238.19 (227.64, 248.75) N = 60	233.12 (220.45, 245.79) N = 40
MMA (ng/g)	2.72 (0.93, 4.51) N = 9	3.01 (2.30, 3.72) N = 8	2.86 (1.33, 4.39) N = 6	**6.84*** (5.69, 7.99) N = 7	4.59 (3.15, 6.04) N = 8	3.60 (2.00, 5.21) N = 8	0.0002	3.90 (3.26, 4.53) N = 46	4.34** (3.56, 5.12) N = 29
DMA (ng/g)	6.94 (2.88, 11.00) N = 9	7.54 (5.59, 9.49) N = 8	6.95 (2.70, 11.20) N = 6	**12.22*** (7.78, 16.65) N = 7	7.34 (5.36, 9.32) N = 8	7.66 (5.91, 9.41) N = 8	0.0808	8.04 (6.87, 9.21) N = 46	8.49 (7.01, 9.98) N = 29
iAs (ng/g)	88.50 (12.96, 164.05) N = 9	46.24 (27.10, 65.39) N = 9	62.10 (38.01, 86.19) N = 6	134.97 (7.20, 262.74) N = 7	51.86 (32.63, 71.08) N = 8	58.01 (17.80, 98.22) N = 8	0.1945	72.53 (50.42, 94.64) N = 47	71.62 (43.59, 99.64) N = 30
Total Arsenic (ng/g)	104.45 (33.97, 175.93) N = 10	81.34 (27.97, 134.72) N = 10	93.14 (33.71, 152.26) N = 10	108.35 (39.46, 177.23) N = 10	59.09 (41.75, 76.43) N = 10	75.90 (37.67, 114.12) N = 10	0.7032	87.05 (67.69, 106.40) N = 60	85.48 (61.79, 109.17) N = 40

Cancer groups including comorbid diabetes (n = 4; 2 bladder cancer, 1 skin cancer, 1 kidney cancer), only 1 skin cancer and 1 bladder cancer for speciation data.

*Significantly different when compared to the healthy reference group (α=0.05) using Dunnett's pairwise comparison test. Note: 95% confidence intervals are only for the individual means per group, while Dunnett's test is for the difference between the means.

**When comorbid diabetes cases dropped, the t-test is significant (mean 4.55, p=0.0248).

Figure 2. Toenail Selenium Concentration by Disease Group



N = 10 per group. Error bars indicate 95% confidence intervals.

4.4 Arsenic Speciation Regression Models for Cancer and Diabetes

An odds ratio estimate for the association of %MMA with type II diabetes was generated using exact logistic regression, given the small sample size. %MMA was significantly associated with having diabetes: for every 1% increase in the proportion of MMA, there was an 89% increase in the odds of having type II diabetes (Table 10). Including the two cases with both diabetes and cancer caused the odds ratio to change from 1.89 to 1.41, becoming not statistically significant ($p = 0.0837$, Table 10).

Table 10. Diabetes Models for %MMA Exposure Using Exact Logistic Regression

	n	OR	P-Value	95% CI
Diabetes	17	1.89	0.0319	(1.03, 4.78)
***Including Cancer**	19	1.41	0.0837	(0.97, 2.38)

*Includes the two comorbid cancer cases

For the "all cancers combined" analysis, there was sufficient sample size to conduct regular logistic regression. Starting with a full model, %MMA and BMI were the only remaining significant variables after backwards stepwise elimination. Age, vigorous physical activity frequency, and ever smoked were sequentially added back in and

compared, with the final model consisting of %MMA, BMI, age, and ever smoked (Table 11).

For every 1% increase in MMA (proportion of total Arsenic) there was a significant increase (114%) in the odds of ever having had primary skin, lung, bladder, or kidney cancer (OR = 2.14), when BMI, age and smoking were held constant (Table 11). Including cases with comorbid diabetes decreased the odds ratio from 2.14 to 1.85 (Table 11). No interaction terms tested were found to be significant in the final models.

Table 11. Crude and Adjusted Combined Cancer Models for %MMA Exposure

	Unadjusted			**Model 1***			**Model 2****		
	OR	P value	95% CI	OR	P value	95% CI	OR	P value	95% CI
%MMA									
Including comorbid diabetes (n=38)*	**1.54**	0.014	(1.09, 2.18)	**1.85**	0.038	(1.03, 3.30)			
Excluding comorbid diabetes (n=36)**	**1.70**	0.010	(1.13, 2.56)				**2.14**	0.036	(1.05, 4.37)
BMI				**0.76**	0.022	(0.60, 0.96)	**0.79**	0.045	(0.632, 0.995)
Age				1.17	0.223	(0.91, 1.50)	1.21	0.189	(0.91, 1.62)
Ever Smoked				15.25	0.075	(0.76, 305.61)	10.24	0.135	(0.49, 215.82)

*Includes the 2 cancer cases with comorbid diabetes.

**Excludes the 2 cancer cases with comorbid diabetes.

Table 12. Cancer Subgroup Analyses for %MMA Exposure using Exact Logistic Regression Models

	n	OR	P-Value	95% CI
Skin Cancer	17	1.68	0.0390	(1.02, 3.23)
***Excluding Diabetes**	16	1.61	0.0645	(0.98, 3.08)
Bladder Cancer	15	1.31	0.2665	(0.84, 2.27)
***Excluding Diabetes**	14	2.11	0.0679	(0.97, 8.63)
Lung Cancer	16	1.46	0.0288	(1.03, 2.36)
Kidney Cancer	17	3.59	0.0016	(1.27, 27.12)

*Excluding comorbid diabetes cases

Subgroup analyses were performed with and without the comorbid diabetes cases. %MMA was significantly associated with an increase in the odds of skin cancer (OR = 1.68), lung cancer (OR = 1.46), and kidney cancer (OR = 3.59), but the association with bladder cancer (OR = 1.31) was not statistically significant (Table 12). In the skin cancer group, including the case with comorbid diabetes did not cause much of a change in the odds ratio (1.68 down to 1.61), though it did cause it to become not statistically significant (p = 0.0645, Table 12). Conversely, including the case with comorbid diabetes in the bladder cancer analysis caused a large increase in the odds ratio (1.31 to 2.11), and brought the estimate closer to significance (p = 0.0679).

4.5 Regression Models for Heterogeneity in Speciation

Backwards stepwise regression for %MMA (as the dependent variable) left cancer history and province in the model as significant predictors. Sleep time and selenium (isotope 82) were close to the significance cut off and were added back into the model, but the best model included only cancer history and province as predictor variables (Model 1, Table 13). Given this model, living in PEI compared to Nova Scotia increases an individual's proportion of MMA by 6.95%, when cancer history is held constant. We are 95% confident that the true % increase in proportion of toenail MMA for the population, for living in PEI compared to Nova Scotia is between 3.86 and 10.04. Having ever had cancer increases an individual's proportion of MMA by 2.39% compared to those with no cancer, when province is held constant. We are 95% confident that the true % increase in proportion of toenail MMA for the population, in patients with cancer history (compared to those without) is between 0.35 and 4.44.

In terms of the distribution of the correlated variables by province, PEI had lower sleep time (N = 6), and higher %MMA (N = 4) (Table 14). Considering these differences, and the fact that all healthy individuals were from Nova Scotia, we made a second model excluding province. Excluding province, the final model included cancer history, sleep time, and toenail selenium concentration as the best predictive variables (Model 2, Table 13). In this model, having ever had cancer increases an individual's proportion of MMA by 2.08% compared to those with no cancer, when sleep time and selenium concentration are held constant. We are 95% confident that the true % increase in proportion of toenail MMA for the population, in patients with cancer history (compared to those without) is between 0.14% and 4.02%.

With every 1 hour increase in sleep time, the proportion of toenail MMA decreases by 0.78% (for 1 min, 0.013% decrease), when cancer history and selenium concentration are held constant. We are 95% confident that the true % decrease in proportion of toenail MMA for the population, for every additional hour of sleep is between 0.07% and 1.50%.

Adding selenium to the model, even though statistically insignificant, changed the odds ratio for cancer slightly, from 2.22 to 2.08. This addition also increased the adjusted R^2, and decreased the AIC. Thus, it was included in the final model.

Table 13. Regression Models for Influences on %MMA

	Model 1*			**Model 2****		
	β	P value	95% CI	β	P value	95% CI
Province						
Nova Scotia	Ref.					
New Brunswick	-0.10	0.928	(-2.33, 2.13)			
Newfoundland	1.27	0.339	(-1.38, 3.91)			
Prince Edward Island	**6.95**	**<0.001**	**(3.86, 10.04)**			
Cancer History						
No	Ref.			Ref.		
Yes	**2.39**	**0.023**	**(0.35, 4.44)**	**2.08**	**0.036**	**(0.14, 4.02)**
Sleep Time (min)				**-0.013**	**0.033**	**(-0.025, -0.001)**
Selenium (ng/g)				-0.019	0.177	(-0.048, 0.009)

N=45 for both models
*Model 1 includes Province and Cancer history, Adj. R2 = 0.3598
**Model 2 includes Cancer history, Sleep time, and selenium Adj. R2 = 0.1827

Table 14. Mean Sleep Time and %MMA by Province

	NS	NB	NL	PEI	P Value
Sleep Time (min)	451.25 (416.21 – 486.29) N=20	462.86 (429.24 – 496.47) N=21	428.18 (375.56 – 480.81) N=11	**345.0*** (268.53 – 421.47) N=6	0.0107
%MMA	4.35 (2.97 – 5.73) N=17	5.67 (4.35 – 6.98) N=18	5.97 (2.95 – 8.99) N=6	**11.65**** (4.14 – 19.16) N=4	0.0007

*significantly different than all using pairwise comparison with bonferroni correction
**significantly different than NS and NB, but not NL, using pairwise comparison with bonferroni correction

Chapter 5. Discussion

5.1 Digestion vs Extraction for Optimal Measurement of Metabolic Species

By using digestion rather than extraction to measure arsenic species, this study was able to achieve an average extraction efficiency of 97.6%, which is much higher than previously reported extraction efficiencies using toenails (35,37). The major drawback to this method is that it cannot distinguish between the forms of inorganic arsenic (As^{III} and As^{V}) due to the oxidative conditions of digestion. However, by forgoing this differentiation we were able to accurately measure the other species found in toenails. In addition to accurate measurement, by using a total inorganic measure we could shorten the time required for speciation analysis by a significant margin. The total run time per sample for Ellingson et al. was 15 minutes (39). By modifying this method and using a summary inorganic measure, we were able to achieve a run time of 6.2 minutes per sample. This considerable time saving also translates to cost savings, as high purity argon and helium gas are required throughout the run.

5.2 Cancer, Diabetes, and Comorbidity Effects

Though the diabetes group had the highest average %DMA as we hypothesized, it was not much higher than the cancer groups, and the difference was not statistically significant (Table 8). In fact, rather than showing the expected "opposite" association compared to cancer, the results showed a similar association, i.e., a higher proportion of monomethylated species associated with increased odds of both cancer and type II diabetes. This result differs from previous findings that diabetes is associated with a lower percentage of MMA in urine (24). This difference in findings could arise from using a different biomarker, or perhaps a different population. The relationship between arsenic speciation patterns in cancer and diabetes will need to be investigated further in future studies.

Including the comorbid cases in both the cancer and diabetes models had some impact on the odds ratios and p-values (Tables 10 - 12). This could be purely a result of chance and the small sample size per group, or could reflect an interaction or effect modification. Having speciation data for only two cases with both diseases makes it impossible to tell from this pilot study, however it would be interesting to investigate further in a subsequent study. Including a full comorbid group would address this

question and allow for more external validity in a future study, as the prevalence of type II diabetes is quite high, and is a common comorbidity in cancer cases.

5.3 Total Arsenic vs Arsenic Speciation

When designing this pilot study, we wanted to get similar arsenic exposure between cases and controls to tease apart the effects of arsenic exposure vs arsenic speciation. We managed to achieve that as the groups did not significantly differ in regards to their total arsenic concentration. No association between total arsenic and cancer or diabetes was found. Instead, we found an association with arsenic speciation, specifically %MMA. This is an interesting finding with serious potential implications. Several studies have found that arsenic is related to multiple forms of cancer at moderate to high exposure levels, but at lower levels, the way the body processes and metabolizes the arsenic could be more important to health than the total amount. It is even possible that there could be an association between arsenic speciation and other cancers previously thought to be unrelated to arsenic exposure. If this were the case, it could call for a re-evaluation of arsenic exposure limits and maximum allowable concentrations in drinking water. However, much more work on a larger scale is needed to inform appropriate policy.

5.4 Impact

This pilot study was the first step toward a better understanding of how arsenic affects human health, and has provided a strong foundation for subsequent larger scale studies to completely characterize the interrelated roles of environmental arsenic exposure, metabolism and body burden (arsenic speciation), lifestyle factors, genetic factors, disease risk, and prognosis. By identifying how potential risk factors and preventative factors interact, not only with arsenic exposure, but with individuals' arsenic speciation/metabolism and body burden, future research has the potential to significantly advance our understanding of arsenic carcinogenicity, and thus allow us to help develop and inform targeted population-level prevention strategies.

The ultimate goal of this research is to contribute to improving health for Atlantic Canadian residents, as well as the millions of people affected by environmental arsenic exposure around the world. Understanding how arsenic causes disease is the first step toward developing and implementing effective intervention programs. This information

can help lead to the improvement of our ability to predict diseases in high-risk sub-populations, and aid in the development of target intervention strategies to prevent the adverse effects of chronic exposure to environmental arsenic. By determining the associations between arsenic speciation profiles and chronic diseases using toenail biomarkers as an indicator of long-term arsenic exposure, this preliminary research has the potential to provide a strong foundation to help generate new hypotheses and guide future research in the exploration of arsenic pathogenesis. The information gained here will add to the growing body of literature to help elucidate the toxic and carcinogenic mechanisms of this global health hazard.

5.5 Strengths and Limitations

The major strengths of this study lie in its novelty: This was the first study to research various cancers and diabetes using toenail arsenic speciation analysis, a potentially important, more relevant biomarker for disease. The current study has improved upon previous speciation research by: 1) extending the scope of arsenic-related cancers being studied, 2) providing a direct and detailed comparison of arsenic speciation profiles between cancer and diabetes, 3) using a more appropriate, longer-term biomarker (toenails) for studying chronic disease associated with long-term arsenic exposure, and 4) providing a detailed arsenic speciation profile (rather than just total arsenic), along with other toxic metal co-contaminants. Additionally, the laboratory portion of the study was able to streamline toenail speciation analysis with far better recoveries than previously reported by opting for a summary inorganic measure rather than differentiating between As^{III} and As^{V}.

The limitations of this study include: 1) The sample size was small, and thus the statistical power was limited, 2) The data were cross-sectional, so we cannot infer causality, 3) The Atlantic PATH cohort is a convenience sample, so selection bias may be an issue. Additionally, the sampling procedure used in this study was province-biased, resulting in all healthy individuals being from Nova Scotia, 4) The questionnaire variables were self-reported and less reliable than objectively measured variables, 5) Given the limited information in the health questionnaire, we were unable to specify the type of skin cancer participants had, 6) With such low survival rates for individuals diagnosed with lung cancer, the sample used in this study may be unrepresentative of typical lung cancer cases, and 7) There was a paucity of arsenic exposure data. We

have assumed that the majority of exposure was from drinking water, though we cannot be certain that this was the case.

5.6 Conclusions and Future Research

Preliminary speciation results indicate that on average, healthy individuals have the lowest percentage MMA, the lowest primary methylation index (PMI), and the highest secondary methylation index (SMI), agreeing with our original hypotheses. Additionally, total metal analysis indicated that selenium was highest in the healthy group. This is interesting as previous work has indicated that selenium can negate arsenic toxicity (33).

Based on these data, we would like to move forward with a larger scale follow-up project, including more participants from each group, and more arsenic-related cancers, especially breast cancer, in which arsenic speciation has been implicated in urinary studies (7). Additionally, it would be interesting to investigate a few cancers that have not been previously associate with arsenic exposure.

The numbers of samples available for analysis from eligible participants at Atlantic PATH have been identified. In addition to the samples already analyzed, toenail samples are available for another 591 skin cancer cases, 19 lung cancer cases, 25 bladder cancer cases, 33 kidney cancer cases, 399 breast cancer cases, and 1472 diabetes cases. In addition to the baseline data, the Atlantic PATH database will continue to grow with regular follow up (every 1 - 2 years), and as such there will be new incident cancer cases to work with, yielding stronger evidence regarding the association between arsenic speciation and cancer.

Given the calculated minimum sample size numbers in Table 15 (Appendix C), we propose to analyze the remainder of the toenail samples available for lung (19 samples), bladder (25 samples), and kidney cancer (33 samples), in addition to another 35 skin cancer, 45 breast cancer, 35 type II diabetes, and 35 healthy samples. Overall, this follow-up work would analyze another 227 samples. Given the minimum numbers calculated in Table 15, these additional samples should allow us enough power to detect significant results for hypothesized associations.

In conclusion, these preliminary speciation results support our central hypothesis that arsenic speciation differs between healthy people and those with chronic arsenic-related diseases. This study provides the first data of its kind using toenails, and

provides solid evidence that this research is both feasible and important, and that further investigation is warranted.

Bibliography

(1) Yu RC, Hsu KH, Chen CJ, Froines JR. Arsenic methylation capacity and skin cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2000;9(11):1259.

(2) Wang W, Cheng S, Zhang D. Association of inorganic arsenic exposure with liver cancer mortality: A meta-analysis. Environ Res 2014;135:120.

(3) Steinmaus N, C., Bates A, M., Yuan L, Y., Kalman E, D., Atallah K, R., Rey H, O., et al. Arsenic Methylation and Bladder Cancer Risk in Case– Control Studies in Argentina and the United States. Journal of Occupational and Environmental Medicine 2006;48(5):478-488.

(4) Ferreccio C, Smith AH, Durán V, Barlaro T, Benítez H, Valdés R, et al. Case-Control Study of Arsenic in Drinking Water and Kidney Cancer in Uniquely Exposed Northern Chile. Am J Epidemiol 2013;178(5):813-818.

(5) Benbrahim-Tallaa L, Waalkes MP. Inorganic Arsenic and Human Prostate Cancer. Environ Health Perspect 2008;116(2):158-164.

(6) Hopenhayn-Rich C, Biggs M, Smith A. Lung and kidney cancer mortality associated with arsenic in drinking water in Cordoba, Argentina. Int J Epidemiol 1998;27(4):561-569.

(7) López-Carrillo L, Hernández-Ramírez RU, Gandolfi AJ, Ornelas-Aguirre J, Torres-Sánchez L, Cebrian ME. Arsenic methylation capacity is associated with breast cancer in northern Mexico. Toxicol Appl Pharmacol 2014;280(1):53-59.

(8) International Agency for Research on Cancer. Some drinking water disinfectants and contaminants, including arsenic. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans 2004;84:512.

(9) Agency for Toxic Substances and Disease Registry. Priority list of hazardous substances. 2015; Available at: http://www.atsdr.cdc.gov/spl/. Accessed November 30, 2015.

(10) Hughes MF, Beck BD, Chen Y, Lewis AS, Thomas DJ. Arsenic Exposure and Toxicology: A Historical Perspective. Toxicological Sciences 2011;123(2):305-332.

(11) Health Canada. Guidelines for Canadian Drinking Water Quality: Guideline Technical Document - Arsenic. 2006.

(12) van Dijk JC, Amy GL, Bakker SA, van Halem D. Arsenic in drinking water: a worldwide water quality concern for water supply companies. Drinking Water Engineering and Science 2009;2(1):29-34.

(13) Smith AH, Lingas EO, Rahman M. Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. Bull World Health Organ 2000;78(9):1093.

(14) Zheng Y, Ayotte JD. At the crossroads: Hazard assessment and reduction of health risks from arsenic in private well waters of the northeastern United States and Atlantic Canada. Sci Total Environ 2015;505:1237-1247.

(15) Goodwin TA, Parker L, Fisher BE, Dummer TJB. Toenails, tap water and you: the arsenic connection. Mineral Resources Branch, Report of Activities 2009; Nova Scotia Department of Natural Resources 2010;1:41.

(16) Goodwin TA, McIsaac EM, Friske PWB. The north american soil geochemical landscapes project: report on the 2008 sampling program and preliminary 2007 results. Mineral Resources Branch, Report of Activities 2008; Nova Scotia Department of Natural Resources 2009;1:45.

(17) McGuigan CF, Hamula CLA, Huang S, Gabos S, Le XC. A review on arsenic concentrations in Canadian drinking water. Env Rev 2010;18:291-307.

(18) Chappells H, Campbell N, Drage J, Fernandez CV, Parker L, Dummer TJB. Understanding the translation of scientific knowledge about arsenic risk exposure among private well water users in Nova Scotia. Sci Total Environ 2015;505:1259.

(19) Nova Scotia Environment. Tips for testing your well water. 2014; Available at: http://www.novascotia.ca/nse/water/welltips.asp. Accessed 04/16, 2017.

(20) Dummer TJB, Yu ZM, Nauta L, Murimboh JD, Parker L. Geostatistical modelling of arsenic in drinking water wells and related toenail arsenic concentrations across Nova Scotia, Canada. Sci Total Environ 2015;505:1248-1258.

(21) Canadian Cancer Society. Canadian Cancer Statistics 2015. 2015.

(22) Canadian Diabetes Association. Diabetes Statistics in Canada. Available at: http://www.diabetes.ca/how-you-can-help/advocate/why-federal-leadership-is-essential/diabetes-statistics-in-canada#_ftn1. Accessed June 30, 2016.

(23) Chen Y, Su H, Guo Y, Hsueh Y, Smith T, Ryan L, et al. Arsenic methylation and bladder cancer risk in Taiwan. Cancer Causes Control 2003;14(4):303-310.

(24) Kuo C, Howard BV, Umans JG, Gribble MO, Best LG, Francesconi KA, et al. Arsenic exposure, arsenic metabolism, and incident diabetes in the strong heart study. Diabetes Care 2015;38(4):620.

(25) Martinez VD, Vucic EA, Becker-Santos D, Gil L, Lam WL. Arsenic Exposure and the Induction of Human Cancers. Journal of Toxicology 2011;2011.

(26) Cui J, Shi J, Jiang G, Jing C. Arsenic levels and speciation from ingestion exposures to biomarkers in Shanxi, China: implications for human health. Environ Sci Technol 2013;47(10):5419.

(27) Tseng C-, E-. A review on environmental factors regulating arsenic methylation in humans. Toxicol Appl Pharmacol 2009;235(3).

(28) Ab Razak NH, Praveena SM, Hashim Z. Toenail as a biomarker of heavy metal exposure via drinking water: a systematic review. Rev Environ Health 2015;30(1):1-7.

(29) He K. Trace elements in nails as biomarkers in clinical research. Eur J Clin Invest 2011;41(1):98.

(30) McCarty KM, Chen Y, Quamruzzaman Q, Rahman M, Mahiuddin G, Hsueh Y, et al. Arsenic Methylation, GSTT1, GSTM1, GSTP1 Polymorphisms, and Skin Lesions. Environ Health Perspect 2007;115(3):341-345.

(31) Normandin L, Ayotte P, Levallois P, Ibanez Y, Marilène Courteau, Kennedy G, et al. Biomarkers of arsenic exposure and effects in a Canadian rural population exposed through groundwater consumption. Journal of Exposure Science and Environmental Epidemiology 2013;24(2):127.

(32) Gagnon F, Lampron-Goulet E, Normandin L, Langlois M. Measurements of Arsenic in the Urine and Nails of Individuals Exposed to Low Concentrations of Arsenic in Drinking Water From Private Wells in a Rural Region of Québec, Canada. J Environ Health 2016;78(6):76.

(33) Chou S, Harper C, Ingerman L, Llados F, Colman J, Chappell J, et al. Toxicological profile for Arsenic. Agency for Toxic Substances and Disease Registry 2007:500.

(34) Vasken Aposhian H, Zakharyan RA, Avram MD, Sampayo-Reyes A, Wollenberg ML. A review of the enzymology of arsenic metabolism and a new potential role of hydrogen peroxide in the detoxication of the trivalent arsenic species. Toxicol Appl Pharmacol 2004;198(3):327-335.

(35) Button M, Jenkin GRT, Harrington CF, Watts MJ. Human toenails as a biomarker of exposure to elevated environmental arsenic. Journal of Environmental Monitoring 2009;11(3):610-617.

(36) Zorimar Rivera-Núñez, Jaymie RM, John DM, Melissa JS, Jerome ON. Urinary arsenic species, toenail arsenic, and arsenic intake estimates in a Michigan population with low levels of arsenic in drinking water. Journal of Exposure Science and Environmental Epidemiology 2011;22(2):182.

(37) Mandal BK, Ogra Y, Suzuki KT. Speciation of arsenic in human nail and hair from arsenic-affected area by HPLC-inductively coupled argon plasma mass spectrometry. Toxicol Appl Pharmacol 2003;189(2):73-83.

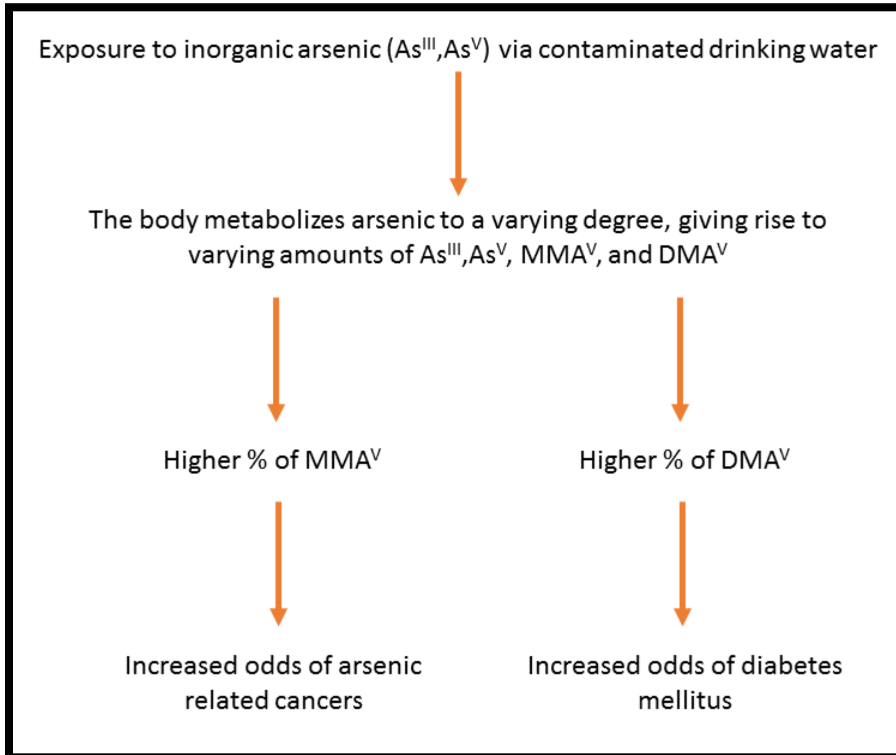
(38) Slotnick M, Meliker J, Avruskin G, Ghosh D, Nriagu J. Toenails as a Biomarker of Inorganic Arsenic Intake From Drinking Water and Foods. Journal of Toxicology and Environmental Health, Part A; Current Issues 2007;70(2):148-158.

(39) Ellingson D, Zywicki R, Sullivan D. Analytical method for the determination of various arsenic species in rice, rice food products, apple juice, and other juices by ion chromatography-inductively coupled plasma/mass spectrometry. J AOAC Int 2014;97(6):1670.

(40) Mehta CR, Patel NR. Exact logistic regression: Theory and examples. Stat Med 1995;14(19):2143-2160.

Appendix A: Hypothesis Diagram

Figure 3. Arsenic Speciation Hypothesis Diagram



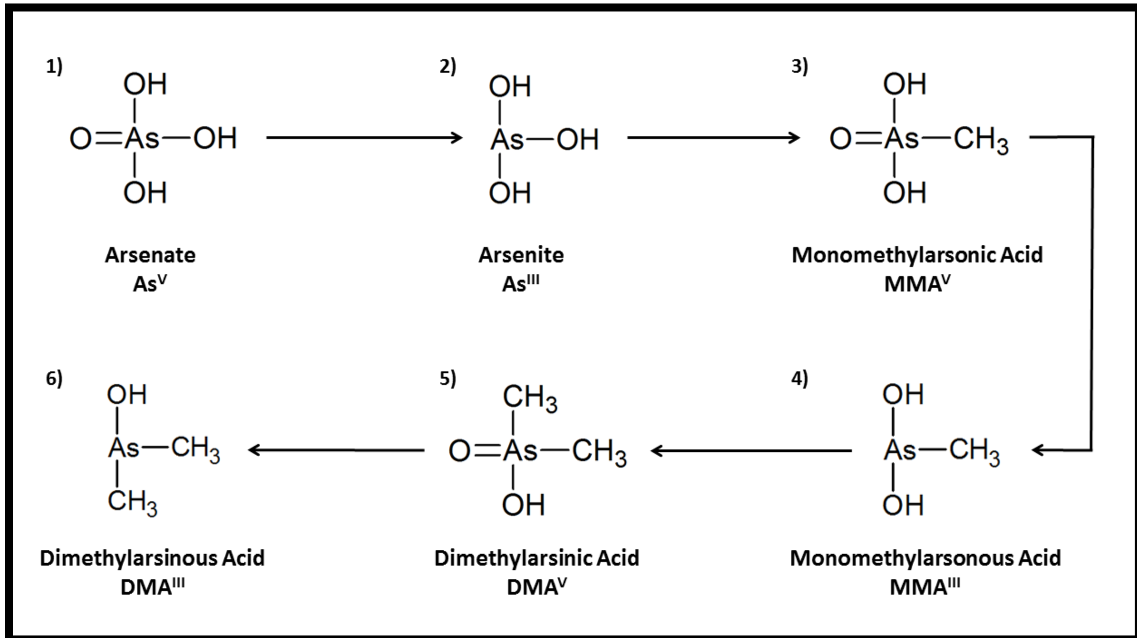
Arsenic speciation is hypothesized to mediate the effect of arsenic exposure on the development of arsenic-related diseases. Specifically, a higher proportion of MMA will increase the odds of having cancer, and a higher proportion of DMA will increase the odds of having diabetes.

*Conventional Studies: Chronic Arsenic exposure → Disease Outcome

*Our Hypothesis: Chronic arsenic exposure → Metabolism and Speciation → Disease Outcome

Appendix B: Arsenic Metabolism Diagram

Figure 4. Typical Arsenic Metabolism Pathway in Mammals



Pathway modified from Aphosian et al. 2004 (34). Individual differences in methylation capacity lead to incomplete transformation, and a mixture of inorganic and mono and dimethylated species are excreted (27).

Appendix C: Sample Size Calculations for Follow-up Study

Table 15. Sample Size Calculations per disease Group for Select Arsenic-related Variables

	Skin	Lung	Bladder	Kidney	Combined Cancer	Diabetes
%MMA	15	12	11	5	10	15
PMI	29	22	28	8	22	31
SMI	694	23	91	24	52	103
82Se	57	6	18	313	32	47
%iAs	33	36	117	17	34	25

All calculations were performed using 95% CI, 80% power, and 1:1 ratio for comparison of two means (each cancer group or combined cancers to healthy). Number were calculation using the group means and standard deviations from the preliminary data (excluding comorbid cancer/diabetes cases) in the equation given below.

$$n_1 = \frac{(\sigma_1^2 + \sigma_2^2/k)\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2}{\Delta^2}$$

$$n_2 = \frac{(k * \sigma_1^2 + \sigma_2^2)\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2}{\Delta^2}$$

N1= sample size for group 1

N2= sample size for group 2

σ1=SD group 1

σ2=SD group 2

Δ=difference in group means

K= ratio=n2/n1

$Z_{1-\alpha/2}$ = two-sided Z value (e.g., / Z=1/96 for 95% CI)

$Z_{1-\beta}$=power

Ref: Bernard Rosner. Fundamentals of Biostatistics (5th edition) (based on equation 8.27).