Arsenic Exposure from Drinking Water and Urinary Metabolomics: Associations and Long-Term Reproducibility in Bangladesh Adults

Fen Wu,1,2 Liang Chi,3 Hongyu Ru,3 Faruque Parvez,4 Vesna Slavkovich,4 Mahbub Eunus,5 Alauddin Ahmed,5 Tariqul Islam,5 Muhammad Rakibuz-Zaman,3 Rabiul Hasan,5 Golam Sarwar,3 Joseph H. Graziano,4 Habibul Ahsan,6 Kun Lu,3 and Yu Chen1,2

1Department of Population Health, New York University School of Medicine, New York, New York, USA
2Department of Environmental Medicine, New York University School of Medicine, New York, New York, USA
3Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
4Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, USA
5U-Chicago Research Bangladesh, Ltd., Dhaka, Bangladesh
6Department of Health Studies, Center for Cancer Epidemiology and Prevention, University of Chicago, Chicago, Illinois, USA

BACKGROUND: Chronic exposure to inorganic arsenic from drinking water has been associated with a host of cancer and noncancer diseases. The application of metabolomics in epidemiologic studies may allow researchers to identify biomarkers associated with arsenic exposure and its health effects.

OBJECTIVE: Our goal was to evaluate the long-term reproducibility of urinary metabolites and associations between reproducible metabolites and arsenic exposure.

METHODS: We studied samples and data from 112 nonsmoking participants (58 men and 54 women) who were free of any major chronic diseases and who were enrolled in the Health Effects of Arsenic Longitudinal Study (HEALS), a large prospective cohort study in Bangladesh. Using a global gas chromatography–mass spectrometry platform, we measured metabolites in their urine samples, which were collected at baseline and again 2 y apart, and estimated intra-class correlation coefficients (ICCs). Linear regression was used to assess the association between arsenic exposure at baseline and metabolite levels in baseline urine samples.

RESULTS: We identified 2,519 molecular features that were present in all 224 urine samples from the 112 participants, of which 301 had an ICC of ≥0.60. Of the 301 molecular features, water arsenic was significantly related to 31 molecular features and urinary arsenic was significantly related to 74 molecular features after adjusting for multiple comparisons. Six metabolites with a confirmed identity were identified from the 82 molecular features that were significantly associated with either water arsenic or urinary arsenic after adjustment for multiple comparisons.

CONCLUSIONS: Our study identified urinary metabolites with long-term reproducibility that were associated with arsenic exposure. The data established the feasibility of using metabolomics in future larger studies.

Introduction

Inorganic arsenic (iAs) occurs naturally in groundwater in many parts of the world, affecting millions of people worldwide. Chronic exposure to iAs from drinking water has been associated with a host of human diseases, including cancer and cardiovascular disease (CVD) (Chen et al. 2009). Metabolomics, or metabolite profiling, refers to the systematic analysis of low molecular weight metabolites (the entire set of metabolites constitute the metabolome) in a biological sample that are intermediates or endpoints of metabolism (Wang et al. 2011). Downstream of genomic, transcriptomic, and proteomic perturbations, metabolites represent the most proximal reporters of alterations in the body in response to external stimuli (Lindon et al. 2003). Metabolomics has the potential to help identify the causes of environmentally mediated disease. Emerging evidence indicates metabolic perturbations associated with exposure to environmental contaminants, including welding fumes (Wang et al. 2012), cadmium (Ellis et al. 2012; Gao et al. 2014; Xu et al. 2016), tobacco smoking (Ellis et al. 2012; Hsu et al. 2013), phthalate (Xu et al. 2016), pesticides (Bonvallot et al. 2013), and arsenic (Li et al. 2017; Martin et al. 2015; Zhang et al. 2014). Epidemiologic studies of arsenic exposure from drinking water and metabolomics are limited. A recent study of 246 pregnant Chinese women identified 9 urinary metabolites that could be used to classify the women into different arsenic exposure levels (Li et al. 2017). Another study in diabetes cases and controls from México found associations between arsenic exposure from drinking water and relative amounts of 61 metabolites in urine (Martin et al. 2015). However, additional population-based studies are needed.

In epidemiologic studies, the usual level of a biomarker is of key interest and most likely to be associated with disease risk or exposure. However, a single measurement in time may not be representative of the usual level, thus reducing the power for epidemiologic studies to detect associations with disease (Rosner et al. 1992). Therefore, it is critical to evaluate the long-term reproducibility of new biomarkers before including them in large epidemiologic studies. Temporal reproducibility refers to the consistency of measurements of more than one sample from the same person at different times (Willett and Lenart 1998) and is expressed by the intraclass correlation coefficient (ICC) as the ratio of between-subject variation to total variation (sum of within- and between-subject variation). The closer the ICC is to 1, indicating little within-subject variation relative to the between-subject variation, the better a single measurement of a biomarker is at differentiating the relative ordering of the level among the individuals (Willett and Lenart 1998). Metabolites in serum and plasma were reported to have, on average, moderate reproducibility (ICC median value 0.4–0.5 covering several months to a year (Floegel et al. 2011; Sampson et al. 2013; Townsend et al. 2013). Urine is easy to collect with a large volume and it is largely free...
from interfering proteins or lipids, presenting opportunities for biomarker discovery in epidemiologic studies. However, the long-term reproducibility of urinary metabolites has yet to be evaluated.

We have established the Health Effects of Arsenic Longitudinal Study (HEALS), a well characterized cohort in Bangladesh with >20,000 participants recruited since year 2000. With repeated measures of urinary arsenic at baseline and every follow-up for more than 95% of the participants, we have the unique opportunity to evaluate the long-term reproducibility of urinary metabolites. We conducted a study of urinary metabolite profiling in 124 HEALS participants with iAs exposure at low-to-moderate levels (0.1 to <300 µg/L). We measured the metabolites in their urine samples collected 2 y apart using a global gas chromatography–mass spectrometry (GC-MS) platform and evaluated long-term reproducibility. We also examined the associations of both water arsenic and urinary arsenic at baseline with metabolites with sufficient reproducibility. In a subset of 84 participants for whom urinary arsenic metabolites had been measured, we also assessed the relationship between monomethylarsonic acid (MMA) percentage, an indicator of arsenic methylation capacity, and these reproducible metabolites.

Methods

Subject Selection

The parent study, the HEALS, is an ongoing prospective cohort study designed to investigate the health effects of arsenic exposure from drinking water in ARAHAZAR, Bangladesh. Details of the HEALS have been described previously (Ahsan et al. 2006). Briefly, between October 2000 and May 2002, we recruited 11,746 married adults (original cohort) 18–75 y of age who were primarily drinking water from a local tube well, from a well-defined 25-km² geographical area. During 2006–2008, the cohort was expanded to include an additional 8,287 participants (expansion cohort) following the same methodologies. The overall participation rate was 97%. At baseline, water samples from all 10,971 tube wells in the study area were collected, and trained clinicians collected demographic and lifestyle data using a standardized questionnaire and collected spot urine samples from participants using structured protocols. The cohort is being actively followed up biennially with similar in-person visits (Ahsan et al. 2006). Spot urine samples were collected at follow-up visits. Informed consent was obtained from the study participants, and the study procedures were approved by the ethical committee of the Bangladesh Medical Research Council and the institutional review boards of Columbia University and the University of Chicago.

All participants in the present study used the tube wells as their exclusive source of drinking water for a long period of time before baseline (on average 7.8 y prior to baseline), and they did not switch wells during the follow-up. Also, analyses of time-series samples collected from 20 tube wells monitored for 3 y in the study area showed that the arsenic concentration in well water was relatively stable over time (Cheng et al. 2005). Therefore, baseline water arsenic is an indicator for long-term exposure in our study population. We aimed to include a homogeneous subpopulation with a wide range of iAs exposure in this study; therefore, we excluded smokers and those with any major chronic diseases such as cancer, CVD, and diabetes from overall HEALS participants. Then we randomly selected a total of 124 participants, consisting of 62 male and 62 female nonsmokers 25–45 y of age. We also frequency matched them by sex, age (±5 y), water arsenic levels (±10 µg/L), and cohort memberships (original vs. expansion cohort).

Arsenic Measurement

Details of the methods have been described (Chen et al. 2013). Briefly, total water arsenic concentration was analyzed by high-resolution inductively coupled plasma mass spectrometry with a detection limit of <0.2 µg/L. Total urinary arsenic concentration was measured by graphite furnace atomic absorption, using a Perkin-Elmer Analyst 600 graphite furnace system (Waltham, MA, USA) with a detection limit of 2 µg/L (Nixon et al. 1991). Urinary creatinine was analyzed using a method based on the Jaffe reaction (Slot 1965). In a subset of 84 HEALS participants, urinary arsenic metabolites were measured by high-performance liquid chromatography (HPLC) separation of arsenobetaine (AsB), arsenocholine (AsC), AsV, AsIII, MMA, and dimethylarsinic acid (DMA), followed by detection by inductively coupled plasma mass spectrometry (Reuter et al. 2003). The percentage of MMA was calculated by dividing MMA by the sum of the metabolites as total arsenic after subtracting AsB and AsC (i.e., nontoxic organic As from dietary sources). Based on our data, urinary MMA% does not change much over time, with an ICC of 0.85 (Ahsan et al. 2007).

Metabolite Measurement Using GC-MS

Urinary metabolites were measured as described previously (Gao et al. 2017). Cold methanol (80 µL) was added to 20 µL urine. After vortexing at maximum speed for 1 min, the samples were incubated at 4°C for 20 min and then centrifuged for 10 min at 12,000 rpm. The supernatant was collected and dried in a SpeedVac (Savant SC110A; Thermo Electron), followed by derivatization using methoxyamine-HCl and BSTFA. The derivatized samples were analyzed using an Agilent Technologies 6890N Network GC System/5,973 Mass-Selective Detector (Agilent Technologies) with an Agilent J&W GC column [30 m length; 0.250 mm diameter (narrow bore); film thickness 0.25 µm] (Agilent Technologies) under the following conditions: initial oven temperature was set at 60°C for 2 min, ramped to 320°C by 8°C/min, and then held at 320°C for 10.5 min. Two microliters of sample solution was injected with helium as the carrier gas at a flow rate of 0.8 mL/min. The temperature of the injector, ion source, and MS Quadrupole were set at 275°C, 230°C, and 150°C, respectively. The mass spectrometer was operated in full scan mode from 50 to 600 m/z. The resultant data were processed with XCMS (https://xcmsonline.scripps.edu) for peak picking, alignment, and extraction of peak intensities. We used molecular features to refer fragment ions obtained by mass spectrometry (Alonso et al. 2015; Lu et al. 2014; Smith et al. 2006) that included both the ions that were assigned to specific metabolites and those with unknown identities. Normalization was performed by dividing the peak area of each molecular feature by the sum of peak areas of all molecular features. The molecular features with an ICC of ≥0.60 were selected for metabolite identification by comparing both the MS spectra and retention time with those in the National Institute of Standards and Technology (NIST) Standard Reference Database.

Statistical Analyses

We used PROC VARCOMP and PROC GLM in SAS (version 9.3; SAS Institute Inc.) to estimate the ICCs and their 95% confidence intervals, respectively, for the normalized peak intensity of each molecular feature detected in the two yearly urine samples. The molecular features with an ICC of ≥0.60 were selected for further statistical analyses. We used linear regression models to
estimate the associations of continuous measure of water arsenic, urinary arsenic, and urinary MMA% with each molecular feature adjusting for sex, age, and cohort memberships. Assumptions of linear regression such as normal distribution of residuals, homocedasticity, and colinearity were checked and none was violated. The results with additional adjustment for body mass index (BMI) were similar and are therefore not shown. The threshold for the significance of the association was adjusted for multiple testing by controlling the false-discovery rate (FDR) (Benjamini and Hochberg 1995). A Venn diagram was used to illustrate the overlap of the metabolites that had a significant association with water arsenic, urinary arsenic, or urinary MMA%. For the reproduceable metabolites that were nominally significantly associated with water arsenic or urinary arsenic, we included a heatmap to present multivariable Pearson correlations between these metabolites and arsenic measures with adjustment for sex, age, and cohort memberships, using the heatmap.2 function from the gplots package in R (version 3.4.0; R Core Team). In addition, we computed least squares means of urinary levels of the metabolite L-threonine that was significantly associated with both water arsenic and urinary arsenic after adjustment for multiple comparisons by quartiles of baseline water arsenic and urinary arsenic levels adjusting for sex, age, and cohort memberships.

Results

Characteristics of the Selected Subjects

A total of 12 participants were excluded from the analysis because metabolites were undetectable in both or one of their two yearly urine samples and the ICC could not be calculated. The present study consisted of 58 (51.8%) men and 54 (48.2%) women who were thin with a mean BMI of 20.5 kg/m², low-educated, and exposed to a mean level of <50 μg/L water arsenic at baseline (Table 1). Men and women did not differ appreciably regarding the matching factors of age, cohort memberships, and baseline total water arsenic as well as other variables such as BMI, systolic blood pressure, diastolic blood pressure, baseline total urinary arsenic and total urinary arsenic at the first follow-up 2 y later; however, men had significantly more years of formal education than women (p < 0.01). Participants in the original cohort had significantly lower systolic and diastolic blood pressure compared with participants in the expansion cohort (p < 0.01), but they did not differ by other variables such as age, BMI, baseline total water arsenic, baseline total urinary arsenic, and total urinary arsenic at the first follow-up (Table 2).

### Table 1. Distribution of selected variables by sex.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Men (n = 58)</th>
<th>Women (n = 54)</th>
<th>p-Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>34.4 ± 5.7</td>
<td>34.9 ± 5.5</td>
<td>0.63</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>19.9 ± 2.6</td>
<td>21.1 ± 4.0</td>
<td>0.07</td>
</tr>
<tr>
<td>Education (y)</td>
<td>4.8 ± 4.0</td>
<td>2.7 ± 3.4</td>
<td>0.003</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>117.6 ± 13.1</td>
<td>115.9 ± 15.7</td>
<td>0.52</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>74.6 ± 10.0</td>
<td>77.3 ± 11.1</td>
<td>0.18</td>
</tr>
<tr>
<td>Baseline total water arsenic (µg/L)</td>
<td>47.2 ± 51.7</td>
<td>44.8 ± 51.4</td>
<td>0.81</td>
</tr>
<tr>
<td>Baseline total urinary arsenic (µg/g creatinine)</td>
<td>194.3 ± 170.1</td>
<td>206.7 ± 178.2</td>
<td>0.71</td>
</tr>
<tr>
<td>Follow-up total urinary arsenic (µg/g creatinine)</td>
<td>171.3 ± 131.2</td>
<td>199.6 ± 162.0</td>
<td>0.31</td>
</tr>
<tr>
<td>Cohort (%): original</td>
<td>25 (43.1)</td>
<td>25 (46.3)</td>
<td>0.74</td>
</tr>
<tr>
<td>Expansion</td>
<td>33 (56.9)</td>
<td>29 (53.7)</td>
<td></td>
</tr>
</tbody>
</table>

*p-Values were computed with the chi-square test or analysis of variance.

### Table 2. Distribution of selected variables by cohort.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Original cohort (n = 50)</th>
<th>Expansion cohort (n = 62)</th>
<th>p-Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>34.3 ± 5.8</td>
<td>35.0 ± 5.4</td>
<td>0.50</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>19.8 ± 2.5</td>
<td>21.0 ± 3.8</td>
<td>0.07</td>
</tr>
<tr>
<td>Education (y)</td>
<td>3.8 ± 4.0</td>
<td>3.9 ± 3.7</td>
<td>0.92</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>112.1 ± 14.5</td>
<td>120.5 ± 13.3</td>
<td>0.002</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>73.6 ± 9.4</td>
<td>77.8 ± 11.2</td>
<td>0.04</td>
</tr>
<tr>
<td>Baseline total water arsenic (µg/L)</td>
<td>55.6 ± 59.4</td>
<td>38.4 ± 42.7</td>
<td>0.08</td>
</tr>
<tr>
<td>Baseline total urinary arsenic (µg/g creatinine)</td>
<td>222.9 ± 191.5</td>
<td>182.0 ± 156.4</td>
<td>0.22</td>
</tr>
<tr>
<td>Follow-up total urinary arsenic (µg/g creatinine)</td>
<td>197.6 ± 162.4</td>
<td>174.8 ± 133.4</td>
<td>0.42</td>
</tr>
</tbody>
</table>

*p-Values were computed with the chi-square test or analysis of variance.
total of 16 metabolites had a confirmed identity (Table 3), namely, aminoethanol, β-amino isobutyric acid, citric acid, 1,2-dithiane-4,5-diol, ethanedioic acid, glycine, 3-hydroxyisovaleric acid, uracil, and uric acid. Furthermore, 6 of the 16 metabolites (1,2-dithiane-4,5-diol, l-threonine, phosphoric acid, pyroglutamic acid, (R*,S*)-3,4-dihydroxybutanoic acid, and succinic acid) were significantly associated with either water arsenic or urinary arsenic after adjustment for multiple comparisons. The metabolite

**Table 3.** Nominally significant associations of reproducible molecular features with baseline total water arsenic and baseline total urinary arsenic.

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>ICC (95% CI)</th>
<th>Water arsenic</th>
<th>Raw p-value</th>
<th>FDR p-value</th>
<th>Urinary arsenic</th>
<th>Raw p-value</th>
<th>FDR p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aminoethanol</td>
<td>0.60 (0.46, 0.70)</td>
<td>0.48 (0.06, 0.89)</td>
<td>0.025</td>
<td>0.11</td>
<td>0.50 (0.09, 0.91)</td>
<td>0.018</td>
<td>0.07</td>
</tr>
<tr>
<td>β-Amino isobutyric acid</td>
<td>0.68 (0.57, 0.77)</td>
<td>-0.02 (-0.05, 0.01)</td>
<td>0.211</td>
<td>0.39</td>
<td>-0.03 (-0.06, -0.00)</td>
<td>0.043</td>
<td>0.11</td>
</tr>
<tr>
<td>Citric acid</td>
<td>0.62 (0.49, 0.72)</td>
<td>-0.03 (-0.05, 0.00)</td>
<td>0.051</td>
<td>0.17</td>
<td>-0.03 (-0.05, -0.00)</td>
<td>0.048</td>
<td>0.12</td>
</tr>
<tr>
<td>1,2-Dithiane-4,5-diol</td>
<td>0.70 (0.59, 0.78)</td>
<td>0.35 (0.09, 0.62)</td>
<td>0.010</td>
<td>0.07</td>
<td>0.58 (0.34, 0.83)</td>
<td>&lt;0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Ethanedioic acid</td>
<td>0.67 (0.56, 0.76)</td>
<td>-0.00 (-0.01, 0.00)</td>
<td>0.200</td>
<td>0.38</td>
<td>-0.00 (-0.01, -0.00)</td>
<td>0.049</td>
<td>0.12</td>
</tr>
<tr>
<td>3-Hydroxyisovaleric acid</td>
<td>0.63 (0.51, 0.73)</td>
<td>-0.02 (-0.04, -0.00)</td>
<td>0.022</td>
<td>0.11</td>
<td>-0.02 (-0.04, 0.00)</td>
<td>0.052</td>
<td>0.12</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.60 (0.47, 0.71)</td>
<td>-0.09 (-0.19, 0.01)</td>
<td>0.072</td>
<td>0.20</td>
<td>-0.11 (-0.20, -0.01)</td>
<td>0.031</td>
<td>0.09</td>
</tr>
<tr>
<td>Indole-3-acetic acid</td>
<td>0.61 (0.48, 0.71)</td>
<td>-0.04 (-0.09, 0.00)</td>
<td>0.060</td>
<td>0.17</td>
<td>-0.05 (-0.09, -0.00)</td>
<td>0.035</td>
<td>0.10</td>
</tr>
<tr>
<td>l-Threonine</td>
<td>0.60 (0.46, 0.70)</td>
<td>-0.13 (-0.22, -0.04)</td>
<td>0.006</td>
<td>0.05</td>
<td>-0.15 (-0.24, -0.06)</td>
<td>0.001</td>
<td>0.02</td>
</tr>
<tr>
<td>Phosphoric acid</td>
<td>0.67 (0.55, 0.76)</td>
<td>1.05 (0.13, 1.96)</td>
<td>0.025</td>
<td>0.11</td>
<td>1.25 (0.35, 2.15)</td>
<td>0.007</td>
<td>0.04</td>
</tr>
<tr>
<td>Pyroglutamic acid</td>
<td>0.64 (0.51, 0.73)</td>
<td>-0.47 (-0.93, -0.01)</td>
<td>0.047</td>
<td>0.16</td>
<td>-0.68 (-1.13, -0.23)</td>
<td>0.003</td>
<td>0.03</td>
</tr>
<tr>
<td>(R*,S*)-3,4-Dihydroxybutanoic acid</td>
<td>0.61 (0.48, 0.72)</td>
<td>-0.09 (-0.16, -0.02)</td>
<td>0.016</td>
<td>0.10</td>
<td>-0.09 (-0.17, -0.02)</td>
<td>0.010</td>
<td>0.04</td>
</tr>
<tr>
<td>Serine</td>
<td>0.63 (0.51, 0.73)</td>
<td>-0.18 (-0.37, 0.02)</td>
<td>0.070</td>
<td>0.20</td>
<td>-0.21 (-0.40, -0.01)</td>
<td>0.035</td>
<td>0.10</td>
</tr>
<tr>
<td>Succinic acid</td>
<td>0.63 (0.50, 0.73)</td>
<td>-0.02 (-0.04, -0.01)</td>
<td>0.003</td>
<td>0.04</td>
<td>-0.02 (-0.04, -0.01)</td>
<td>&lt;0.001</td>
<td>0.02</td>
</tr>
<tr>
<td>Uric acid</td>
<td>0.63 (0.50, 0.73)</td>
<td>-0.05 (-0.11, -0.01)</td>
<td>0.127</td>
<td>0.28</td>
<td>-0.07 (-0.13, -0.01)</td>
<td>0.031</td>
<td>0.09</td>
</tr>
<tr>
<td>Total</td>
<td>0.64 (0.51, 0.73)</td>
<td>2.56 (0.39, 4.74)</td>
<td>0.022</td>
<td>0.11</td>
<td>1.27 (-0.93, 3.47)</td>
<td>0.255</td>
<td>0.35</td>
</tr>
</tbody>
</table>

*Coefficient from linear regression model indicates difference in peak intensity of urinary metabolites in relation to per 1-SD increase in water arsenic (51.3 µg/L) and per 1-SD increase in urinary arsenic (173.4 µg/g creatinine), adjusting for sex, age, and cohort memberships.
1-threonine was significantly associated with both water arsenic and urinary arsenic after adjustment for multiple comparisons (Table 3). The adjusted means of urinary levels of 1-threonine according to quartiles of baseline water arsenic and urinary arsenic levels are shown in Figure 2. Overall, 1-threonine levels were inversely related to water arsenic and urinary arsenic in a similar pattern and were significantly lower in the highest quartile compared with the lowest quartile ($p < 0.05$ for water arsenic and $p < 0.01$ for urinary arsenic).

**Discussion**

To investigate the suitability of urinary metabolite profiling for use in searching for biomarkers of arsenic-related health effects, we conducted a study to evaluate the long-term reproducibility of the metabolites using repeated urine samples collected 2 y apart. Our data showed that there are substantial known and unknown metabolites with sufficient reproducibility and strong associations with iAs exposure, presenting future opportunities of biomarker discovery for epidemiologic studies.

We found 301 molecular features (27% of the detected features) with excellent reproducibility over time (ICCs $\geq 0.60$). These ICCs compare favorably with the reproducibility over a several-year period of serum cholesterol (ICC = 0.65) (Shekelle et al. 1981), blood pressure (ICC = 0.60–0.64) (Rosner et al. 1977), blood glucose (ICC = 0.52) (Gordon and Shurtleff 1973), pulse (ICC = 0.49) (Gordon and Shurtleff 1973), and plasma estradiol in postmenopausal women (ICC = 0.68) (Hankinson et al. 1995), all of which are exposures considered to be reasonably well-measured and reliable predictors of disease in epidemiologic studies. More than 10% of the 301 molecular features were also associated with urinary MMA%, a biomarker specific for susceptibility to iAs exposure that has been related to cancers (Chen et al. 2003; Huang et al. 2008; Steinmaus et al. 2010; Yu et al. 2000). Furthermore, 82 of the 301 molecular features were significantly associated with either water arsenic or urinary arsenic after controlling for the influence of sex, age, cohort memberships and multiple comparisons. Although a limited number of metabolites were identified from these reproducible molecular features, possibly because of the intrinsic limitation of GC-MS, our results suggest that within the context of a prospective epidemiologic study, a single urine measurement of certain molecular features may adequately represent their longer-term usual (i.e., at least 2 y) levels, which may serve as intermediate biomarkers linking arsenic exposure and chronic diseases.

Consistent with our finding of reproducible molecular features, a previous study investigated the source of variability of 539 metabolites measured by LC-MS and GC-MS in urine samples and found that a large proportion (81%) of the metabolites had an ICC exceeding 0.50. However, the evaluation of reproducibility was based on 17 male subjects over 2 to 10 d. In our study, reproducibility of molecular features were estimated in urine samples collected 2 y apart and we found 27% molecular features having an ICC over 0.50. The difference in time intervals between sample collections may partially explain the much higher portion of reproducible metabolites in the previous study as compared with that of our study. For many metabolites, the correlation between samples would be expected to be higher when time intervals between sample collections are short. As a result, measures over a few days may not capture the true temporal variability around the “usual” long-term level of these metabolites. Therefore, our study adds to the evidence that certain urinary metabolites may be more relatively stable than others over a longer period of time and have the potential to serve as long-term biomarkers associated with exposures and/or diseases.

We identified several urinary metabolites that were dose-dependently associated with either water arsenic or urinary arsenic, some of which may be of biological significance. For instance, three amino acids (glycine, 1-threonine, and serine) that are involved in one-carbon metabolism—the central pathway that facilitates arsenic methylation and elimination—were inversely related to water or urinary arsenic levels. Glycine and serine participate in the metabolism of methionine as a methyl-group acceptor and as a substrate for cystathionine synthesis, respectively (Benevenga and Harper 1970; Stead et al. 2000). Both animal and human studies (Benevenga and Harper 1970; Fukada et al. 2006; Girard-Globa et al. 1972; Stead et al. 2000; Verheof et al. 2004) have shown that serine can lower homocysteine—a risk factor for CVD. Glycine is synthesized endogenously from serine, threonine, choline, or glyoxylate in the liver and kidney (Wang et al. 2013). Glycine exerts anti-inflammatory and antioxidative effects (McCart and DiNicolantonio 2014; Senthilkumar et al. 2004) and has been shown to reduce plasma insulin, fat mass, and blood pressure in rodents (Alvarado-Vásquez et al. 2003; El Hafidi et al. 2004). Lower glycine concentrations have been associated with several traditional cardiovascular risk.

**Figure 2.** Adjusted means of normalized peak intensity of 1-threonine by quartiles of baseline (A) total water arsenic levels and (B) total urinary arsenic levels. Means were adjusted for sex, age, and cohort memberships.

---

*Environmental Health Perspectives* 017005-5
Two epidemiologic studies have investigated the impact of arsenic exposure from drinking water on metabolite profiles. A recent study of 246 pregnant Chinese women identified nine urinary metabolites that could be used to classify the women into a low (the first tertile of urinary total creatinine-adjusted arsenic) or high (the third tertile of urinary arsenic) exposure category using UPLC/Q-TOF MS (ultra performance liquid chromatography coupled to quadrupole with time-of-flight mass spectrometry) (Li et al. 2017). The identified metabolites were potentially related to endocrine disruption and oxidative stress. Another study of 86 Mexican individuals with exposure to low-to-moderate arsenic levels in drinking water (0.1 to 285 μg/L) reported 61 altered metabolites in urine associated with urinary total unadjusted arsenic using GC- and LC-TOF-MS; these metabolites were associated with amino acid metabolism, carbohydrate/energy metabolism, and vitamin (riboflavin) metabolism (Martin et al. 2015). However, there was no overlap of the identified metabolites between these studies and our study. It should be noted that we evaluated the exposure–metabolite associations only for the metabolites that were relatively stable over time (ICC ≥ 0.60). Taken together, these data point to metabolic disruption by arsenic exposure, though specific metabolites shared across all studies are difficult to identify because of differences in exposure background, population characteristics, and metabolomics platforms. Larger studies are needed to characterize the interplay between arsenic exposure, metabolite profile, and disease outcomes.

Strengths of this study include a sufficient number of subjects for a metabolomics study to assess reproducibility, a wide range of arsenic exposure levels, and the long interval between collections of repeated urine samples for evaluation of long-term reproducibility. This study is among the few to evaluate the long-term reproducibility of metabolomics data in human urine samples from a prospective cohort study. We also acknowledge several limitations. First, GC-MS data contain complexity such as that a simple strategy is to combine these fragments with the reproducible metabolites from all the 2,519 molecular features because our aim was to evaluate the existence of reproducible metabolomics data using urine samples that are readily available from our large parent cohort study. Second, we identified only a small number of metabolites from the reproducible molecular features that prevented us from further pathway analysis. We therefore acknowledge that a complementary platform such as LC-MS should be employed in future studies for comprehensive understanding of metabolic alterations in response to arsenic exposure. In summary, our study identified urinary metabolites with long-term reproducibility that were associated with arsenic exposure using a global GC-MS metabolomics platform. The data established the feasibility of using metabolomics platform in future larger studies to assess alterations in urinary metabolites in relation to arsenic exposure and their associations with the risk of CVD or cancer.

References


