Quantitative Drinking Water Arsenic Concentrations
Measured Using Mobile Phone Photometry and Field Kits

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Environmental Science &amp; Technology Letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>Draft</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Letter</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Haque, Ezazul; Lamont-Doherty Earth Observatory Mailloux, Brian; Barnard College, Dept of Environmental Science DeWolff, Daisy; Barnard College, Environmental Sciences Gilioli, Sabina; Barnard College, Environmental Sciences Kelly, Colette; Barnard College, Environmental Sciences Small, Christopher; Columbia University, Lamont-Doherty Earth Observatory Ahmed, Ershad; University of Dhaka, Ahmed, Kazi; Dhaka University Van Geen, Alexander; Columbia University, Lamont-Doherty Earth Observatory Bostick, Benjamin; Columbia University, Lamont Doherty Earth Observatory</td>
</tr>
</tbody>
</table>
Quantitative Drinking Water Arsenic Concentrations Measured
Using Mobile Phone Photometry and Field Kits

EZAZUL HAQUE\textsuperscript{a,b}, BRIAN J. MAILLOUX\textsuperscript{b}, DAISY DE WOLFF\textsuperscript{b}, SABINA GILIOLI\textsuperscript{b},
COLETTE KELLY\textsuperscript{b}, ERSHAD AHMED\textsuperscript{c}, CHRISTOPHER SMALL\textsuperscript{a}, KAZI MATIN
AHMED\textsuperscript{c}, ALEXANDER VAN GEEN\textsuperscript{a}, BENJAMIN C. BOSTICK\textsuperscript{a}*

\textsuperscript{a} Lamont-Doherty Earth Observatory of Columbia University, Route 9 W, Palisades, New York 10964, United States
\textsuperscript{b} Department of Environmental Sciences, Barnard College, 3009 Broadway, New York, New York 10027, United States
\textsuperscript{c} Department of Geology, Dhaka University, Dhaka 1000, Bangladesh

* Corresponding author: Benjamin C. Bostick
Phone: (+1) 845-365-8659; Fax: (+1) 845-365-8155; E-mail: bostick@ldeo.columbia.edu

Prepared for submission to \textit{Environmental Science and Technology Letters}
Draft Date: August 1, 2016
Photo abstract
Abstract

Arsenic (As) groundwater contamination is common yet spatially heterogeneous within most environments. It is therefore necessary to measure As concentrations to determine whether a water source is safe to drink. Measurement of As in the field involves using a test strip that changes color in the presence of As. These tests are relatively inexpensive, but results are subjective and provide binned categorical data rather than exact determinations of As concentration. The goal of this work was to determine if photos of field kit test strips taken on mobile phone cameras could be used to extract more precise, continuous As concentrations. As concentrations for 376 wells sampled from Araihazar, Bangladesh were analyzed using ICP-MS, field kit and the new mobile phone photo method. Results from the field and lab indicate that normalized RGB color data extracted from images were able to accurately predict As concentrations as measured by ICP-MS, achieving detection limits of 9.2 µg/L, and 21.9 µg/L for the lab and field respectively. This work indicates that mobile phone cameras can be used as an analytical for quantitative measures of As and could change how water samples are analyzed.
Introduction

Arsenic (As) exposure in drinking water can cause cancers of the skin, bladder, and lung, and has been associated with other adverse health effects including skin lesions, reproductive effects, nonmalignant pulmonary disease, cardiovascular disease, and other illnesses\(^1-^5\). About 20 million to 45 million people are exposed to concentrations above the Bangladesh national standard of 50 µg/L and the World Health Organization’s (WHO) guideline value of 10 µg/L, respectively\(^6\). Within As impacted environments, As concentrations can vary in both location and depth. Thus, being able to accurately and quickly determine As concentrations in water is imperative in alleviating the burden of disease and for the protection of public health.

In Bangladesh, labeling wells with their As concentrations reduces exposures and promotes households with a high arsenic well to switch to a low arsenic well\(^7\). The Multiple Cluster Indicator Survey in 2009 showed that in Bangladesh, 44% of well owners did not know the status of As in their wells\(^8\). Methods for determining As concentrations include laboratory and/or field-testing. Laboratory methods of analyses, such as inductively coupled plasma mass spectrometry (ICP-MS) and flow injection hydride generation atomic absorption spectroscopy (FI-HG-AAS), provide high precision measurements but can be costly, take place far from the water source, and require days to weeks for analysis, making it difficult to provide feedback to households, and for households to make informed decisions based on their results. Field methods often involve visually comparing the color of a test strip with a reference chart of colors that vary with As concentrations. These field methods are less expensive and allow for rapid, on-site sample analysis and immediate feedback to the households. Field kits can be effective at discriminating As above and below the Bangladesh drinking water standard of 50 µg/L when performed by trained technicians\(^9-\text{\textsuperscript{11}}\), but don’t provide continuous or discrete As concentrations.
Improved analysis of test kit color changes could enable improved utilization of test results. The use of digital image processing has previously been shown to provide accurate determinations of field kit color changes that can be quantitatively transformed to concentrations in controlled laboratory settings\textsuperscript{12-16}. Digital sensors on current smartphone cameras have sufficient resolution and sensitivity to provide color values from various digital color spaces that are used to create, represent and visualize colors with a minimum number of channels. The most popular color space model is the RGB model, in which each sensor captures the intensity of the light in the red (R), green (G) or blue (B) portion of the spectrum. By extracting these colors, it is possible to digitally analyze the images to objectively quantify color, and thereby concentration in environmental samples.

Here, we present a novel method that transforms the information provided in existing As field test kits, and apply this method in the field, where conditions can deviate considerably from ideal lab settings. Water samples were collected from private tube wells in Bangladesh and immediately analyzed by field kits. The test strip of the field kit was photographed next to a standard color chart using a cell phone. The photograph was analyzed to determine concentration and compared to laboratory measurements. This approach indicates that cell phone photometry can provide an accurate field method of quantitatively measuring groundwater As concentrations.

Methods

Field Sampling

The field site is in the Araihazar upazilla located approximately 20 km east of Dhaka in Bangladesh. The study area is a 25 km$^2$ portion of Araihazar which consists of 61 villages. Over
the past two decades, it has been the site of numerous health and geochemical studies\textsuperscript{2,17-19}. For this study, village health workers (VHW) sampled 376 tube wells that have been part of previous studies\textsuperscript{7,20-23}. The tube wells were sampled and analyzed by both field kits and ICP-MS during this study. During sampling, the VHWs collected the pH, Eh, and surveyed the head of household. From each well, groundwater was collected and stored in 20 mL scintillation vials for laboratory analysis. A separate sample of 50 mL was also collected for field analysis done on site using the EconoQuick test kit (Industrial Test Systems Inc. \texttt{http://www.sensafe.com/}) according to the manufactures protocol and previous studies\textsuperscript{9-11,24} with the addition of a photograph taken of the strip at the end of the test.

**Digital Image Capturing and Processing**

Field photos were taken with a Samsung S Duos-2 (Model no. GT-S7582L) mobile phone, which has a resolution of 5 megapixels yielding photos of 1.3 megabytes JPEG files. Laboratory photos were taken with an iPhone 5S (Model no. ME308LL/A) mobile phone, which has a resolution of 8 megapixels yielding photos of 2.7 megabytes JPEG files. Written directions, translated into Bangla, provided to the VHWs asked that each photo be taken outdoors with indirect sunlight from approximately 60cm at a 45° angle with the sample test strip, an unused test strip, the EQ arsenic test kit concentration color chart, and a color checker card (DGK Color Tools, \texttt{http://www.dgkcolortools.com}), (Figure SI-3 to Figure SI-8).

Digital image processing was conducted using Adobe Photoshop CS6 (Version 13.0 x64, \texttt{http://www.adobe.com/products/photoshop.html}). Each photo was normalized using the Levels Adjustment tool. The goal of the normalization is to minimize the effect of lighting conditions and enable the determination of a consistent set of colors. Levels tool was used to carry out black normalization for the photo. White and gray normalizations were also evaluated and did not
provide reliable results for concentrations below 100 µg/L due to overexposure of lighter colors, making it difficult to discriminate colors between the lower concentrations (data not shown). The Marquee tool was used to delineate the area of an image for color extraction. The average RGB color value of all the pixels in the test strip and the standards of the EQ standard concentration chart were extracted and recorded for analysis.

Diffuse reflectance of the test strips was measured using a Konica Minolta CR-700d Spectrophotometer (Konica Minolta Sensing Americas, Inc. Ramsey, New Jersey, USA) within the visible spectrum ranging from 360 to 750 nm. Absorbance at the three channels of R, G, and B were measured at 700, 550 and 400 nm respectively. Arsenic standards were made using sodium arsenate (Fisher Scientific, Pittsburgh, PA) and concentrations were verified by ICP-MS.

ICP-MS Analysis

Groundwater samples collected in 20 mL scintillation vials were acidified to 1% with high-purity Optima Grade Nitric Acid at Lamont-Doherty Earth Observatory at least 48 hours before analysis. Water samples were diluted 1:10 in a solution spiked with $^{73}$Ge for internal drift correction and analyzed for As by high-resolution inductively coupled plasma mass spectrometry (Thermo Fisher Scientific, Model: Element XR).

Statistical Analysis

Statistical analysis was done using iPython Version 2.7 with the Pandas, Statsmodels, and Scipy packages. Maps were created using ArcGIS (ESRI, http://www.esri.com/software/arcgis). The RGB color value of the EQ standard concentration color chart values at 25 µg/L were subtracted from the other RGB values within each photo to further normalize the data. A multiple linear regression between the normalized RGB color values and the known As concentrations was performed.
Results and Discussion

There is a clear relationship between the visible diffuse reflectance spectrum of field kit test strips and As concentration (Figure 1). The tests strip changes from a white to yellow to a dark red as As concentration increases. Reflectance of test strips decreased with As concentration at the fixed wavelengths of 400, 550, and 700 nm (the center of red, green and blue light respectively), and this decrease was mirrored by parallel but nonlinear decreases in red, green and blue (RGB) color values obtained from images of test strips at the same concentrations (Figure 1). These data indicate that each channel responds differently to concentration and thus is sensitive to different concentration ranges. The blue channel is most sensitive and shows a response that is approximately linear from 0-150 µg/L As, whereas green reflectance is linear over a greater concentration range, and red reflectance is only sensitive to concentrations >500 µg/L. These variations in reflectance are exploited to develop standard curves for As concentrations.

Four independent laboratory standard curves with varying concentrations up to 1280 µg/L were analyzed. Each standard curve was photographed in one image taken in the laboratory under controlled lighting conditions. The images showed consistent decreases in all channels, though blue reflectance changed most significantly. Individual standard curves performed much better, in part because they were obtained under uniform, broad-spectrum lighting and with a single camera in a single photograph (SI figures 3-7).

Over the entire range the individual and combined standard curves had $r^2$ values >0.95 and 0.90 respectively. There is considerable error and bias in the regression due to non-linearity in the fitting. Because blue reflectance decreases linearly to 150 µg/L and then flattens, we focused our analysis on the low concentration range, under which blue reflectance is linear with
concentration (<150 µg/L). This improved the $r^2$ to >0.96 for individual standard curves and to 0.95 for all lab data. The detection limits also improved to 8.2 µg and 9.2 µg/L for individual and aggregated lab calibrations. The strong correlations and low detection limits in the 0-150 µg/L range indicate what is possible under well controlled conditions, and suggest that this method should also accurately measure As concentrations in the field.

In all, 376 wells were sampled for As and analyzed by the field kit. 288 field photos met the criteria of having the reference color card, field test kit chart, and a blank unused test strip next to the sample test strip. Of these, 274 wells had water samples for As which were analyzed by ICP-MS; 51 had As <10 µg/L, 79 had 10 to 50 µg/L, 90 had 50 to 150 µg/L, and 54 had >150 µg/L. The changes in RGB of the field photos were consistent with the lab-based trends. Over the entire concentration range of field samples, the $r^2$ was 0.750 with a detection limit of 61 µg/L (Figure SI-8). The non-linear change in color over the entire range limits the method of quantitation—the regression model is linear and the response in reflectance with concentration is nonlinear. When limiting analysis to the linear range (0-150 µg/L) results were much improved. 220 samples had As <150 µg/L and had an $r^2$ of 0.827 with a detection limit of 22 µg/L (Figure SI-8). In the future, other quantitation methods, such as piecewise regression, factor analysis, or machine learning should improve detection limits and linear ranges.26

There is broad agreement between As concentrations predicted with the test kit photo color data and ICP-MS measured concentrations for field samples (Figure 2). For most samples, their predicted concentration falls near the 1:1 line, but modeled and actual As concentrations differ in many cases. Laboratory-based photos however all fall within 5% of the 1:1 line, indicating that this method accurately quantifies concentrations as low as 10 µg/L in the laboratory under uniform conditions. Variation in field quantification reflects the important
effect of lighting on reflectance spectra (something that likely affects our visual classification of
concentration as well). We attribute the lower correlation in the field samples to this variability
and other potential errors resulting from running the test kit. Nevertheless, the method was able
to predict concentrations without bias or correction that is typical of test strip analyses (Figure
SI-9). Indeed a visual inspection of the images from this unsupervised field trial reveals
considerable inconsistency in angle, focus, distance, lighting and levels of specular reflectance,
all of which affect RGB results\textsuperscript{27, 28}. In subsequent field efforts, it should be possible to better
control image collection conditions and analysis methods, to achieve more uniform white
balance and minimize specular reflectance component of the test strips and calibration cards to
correct RGB values thereby improving method accuracy and precision.

This method transforms an existing field kit, which provides binned categorical data, into
a method that provides continuous concentration data. Continuous data improves exposure
assessment and the ability to track spatial and temporal variability. For example, kriging with
both the ICPMS and the photo predicted data helps to delineate high and low As regions that are
misclassified by binned field kit data and to better define the boundaries between them (Figure
3).

Results from this study highlight the potential for mobile phone photometry to accurately
determine As concentrations in the field. This method could be applied to many other
colorimetric kits. In the future, the method should improve as phones with higher resolution and
dynamic range are introduced. Automation of our method with a mobile phone app could enable
widespread high quality data acquisition by individual stakeholders and citizen scientists.
Acknowledgement

We acknowledge the help of T. Ellis, M. Mozumder, and the village health workers. This study was funded by The Barnard College Summer Research Institute, the National Institute of Environmental Health Sciences (grants ES010349 and ES009089), and National Science Foundation Chemistry grant 1310368. This is Lamont Contribution no. XXXX (the number will be added if accepted).
References


Table 1: Summary of statistical results for both laboratory and field analyses. Slope, intercept, and \( r^2 \) coefficient were calculated from linear regression model of photo predicted versus ICP-MS determined As for concentrations up to 150 \( \mu \text{g/L} \), and for full range with maximum As concentration display to the right. Laboratory and field detection limits (DL) based on predicted values from regression up to 150 \( \mu \text{g/L} \), and full range. Field blanks were classified by ICP-MS determined concentrations below 1 \( \mu \text{g/L} \). If there were less than 5 blanks, the detection limit was not provided.

<table>
<thead>
<tr>
<th>Date</th>
<th>Label</th>
<th>Figure</th>
<th>n</th>
<th>( r^2 )</th>
<th>DL (( \mu \text{g/L} ))</th>
<th>n</th>
<th>( r^2 )</th>
<th>DL (( \mu \text{g/L} ))</th>
<th>Max As (( \mu \text{g/L} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>April 16\textsuperscript{th}, 2015</td>
<td>Lab 1</td>
<td>SI-1</td>
<td>5</td>
<td>1.00</td>
<td>-</td>
<td>9</td>
<td>0.96</td>
<td>-</td>
<td>1000</td>
</tr>
<tr>
<td>April 23\textsuperscript{rd}, 2015</td>
<td>Lab 2</td>
<td>SI-2</td>
<td>13</td>
<td>0.96</td>
<td>-</td>
<td>17</td>
<td>0.99</td>
<td>-</td>
<td>1000</td>
</tr>
<tr>
<td>Oct, 7\textsuperscript{th}, 2015</td>
<td>Lab 3</td>
<td>SI-3</td>
<td>14</td>
<td>0.97</td>
<td>-</td>
<td>24</td>
<td>0.95</td>
<td>-</td>
<td>1280</td>
</tr>
<tr>
<td>Jul 30\textsuperscript{th}, 2016</td>
<td>Lab 4</td>
<td>SI-4</td>
<td>10</td>
<td>0.99</td>
<td>8.2</td>
<td>13</td>
<td>0.97</td>
<td>150</td>
<td>1000</td>
</tr>
<tr>
<td>2015-2016</td>
<td>Lab 1-4</td>
<td>SI-5</td>
<td>42</td>
<td>0.96</td>
<td>9.2</td>
<td>62</td>
<td>0.90</td>
<td>180</td>
<td>1280</td>
</tr>
<tr>
<td>July, 2015</td>
<td>Field</td>
<td>SI-6</td>
<td>220</td>
<td>0.83</td>
<td>22</td>
<td>274</td>
<td>0.75</td>
<td>61</td>
<td>750</td>
</tr>
</tbody>
</table>
Figure 1: Relationship between photo extracted RGB color values (left y-axis) and reflectance (right y-axis) of test strips as a function of As concentration measured by ICP-MS. Lab photos are displayed in solid colors corresponding to their color channels, field photos are shown in gray, and reflectance in black. Reflectance measurements were done for laboratory prepped solutions. A) R Corrected, B) G Corrected, and C) B Corrected.
Figure 2: Linear regression of photo predicted against ICP-MS predicted As concentration for lab samples (red) and field samples (blue).
Figure 3: Contour plots of a) ICPMS determined b) field kit photo predicted and c) field kit strip predicted As concentration expressed in µg/L. The location and As concentrations of wells sampled are displayed in d. Well with depths >50 m are not shown.
Supporting Information

Quantitative Drinking Water Arsenic Concentrations Measured Using Mobile Phone Photometry and Field Kits

EZAZUL HAQUE\textsuperscript{a,b}, BRIAN J. MAILLOUX\textsuperscript{b}, DAISY DE WOLFF\textsuperscript{b}, SABINA GILIOLI\textsuperscript{b}, COLETTE KELLY\textsuperscript{b}, ERSHAD AHMED\textsuperscript{c}, CHRISTOPHER SMALL\textsuperscript{a}, KAZI MATIN AHMED\textsuperscript{c}, ALEXANDER VAN GEEN\textsuperscript{a}, BENJAMIN C. BOSTICK\textsuperscript{a}

\textsuperscript{a} Lamont-Doherty Earth Observatory of Columbia University, Route 9 W, Palisades, New York 10964, United States
\textsuperscript{b} Department of Environmental Sciences, Barnard College, 3009 Broadway, New York, New York 10027, United States
\textsuperscript{c} Department of Geology, Dhaka University, Dhaka 1000, Bangladesh
Figure SI-1: Absorbance spectrum of field kit test strips of a blank and tests used for lab solutions of As concentrations of 10, 50, 150, and 500 µg/L.
Figure SI-2: Relationship between ICP-MS measured As concentration (µg/L) for RGB color values a) not normalized and b) normalized. Field and lab samples are represented by gray and blue points respectively.
Figure SI-3: Summary of experiment done on April 16th, 2015 in laboratory controlled conditions showing the black normalized photograph taken on an iPhone 5S (A), relationship between normalized R, G, B with respect to As concentration (B), inductively coupled plasma mass spectrometry (ICP-MS) determined versus photo predicted As concentration for full range up to 1500 µg/L, and an expanded version of (C) is shown in (D) for As concentrations up to 150 µg/L.
Figure SI-4: Summary of experiment done on April 23rd, 2015 in laboratory controlled conditions showing the black normalized photograph taken on an iPhone 5S (A), relationship between normalized R, G, B with respect to As concentration (B), inductively coupled plasma mass spectrometry (ICP-MS) determined versus photo predicted As concentration for full range up to 1500 µg/L, and an expanded version of (C) is shown in (D) for As concentrations up to 150 µg/L.
Figure SI-5: Summary of experiment done on October 7th, 2015 in laboratory controlled conditions showing the black normalized photograph taken on an iPhone 5S (A), relationship between normalized R, G, B with respect to As concentration (B), inductively coupled plasma mass spectrometry (ICP-MS) determined versus photo predicted As concentration for full range up to 1500 µg/L, and an expanded version of (C) is shown in (D) for As concentrations up to 150 µg/L.
Figure SI-6: Summary of experiment done on October 7\textsuperscript{th}, 2015 in laboratory controlled conditions showing the black normalized photograph taken on an iPhone 5S (A), relationship between normalized R, G, B with respect to As concentration (B), inductively coupled plasma mass spectrometry (ICP-MS) determined versus photo predicted As concentration for full range up to 1500 \(\mu\text{g/L}\), and an expanded version of (C) is shown in (D) for As concentrations up to 150 \(\mu\text{g/L}\).
Figure SI-7: Summary of experiment done in laboratory controlled conditions from Figure SX-SX showing the black normalized photograph taken on an iPhone 5S (A), relationship between normalized R, G, B with respect to As concentration (B), inductively coupled plasma mass spectrometry (ICP-MS) determined versus photo predicted As concentration for full range up to 1500 µg/L, and an expanded version of (C) is shown in (D) for As concentrations up to 150 µg/L.
Figure SI-8: Summary of tests done in field in Araihazar, Bangladesh on July-2015 showing the black normalized photograph taken on an iPhone 5S (A), relationship between normalized R, G, B with respect to As concentration (B), inductively coupled plasma mass spectrometry (ICP-MS) determined versus photo predicted As concentration for full range up to 1500 µg/L, and an expanded version of (C) is shown in (D) for As concentrations up to 150 µg/L.
Figure SI-9: Linear regression ICP-MS measured As concentrations versus strip predicted field results as determined in the field.