Quantitative Drinking Water Arsenic Concentrations in Field Environments Using Mobile Phone Photometry of Field Kits

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

[a] *Lamont-Doherty Earth Observatory of Columbia University, Route 9 W, Palisades, New York 10964, United States*
[b] *Department of Environmental Sciences, Barnard College, 3009 Broadway, New York, New York 10027, United States*
[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

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. Data analysis is most consistent in the laboratory, but can successfully be carried out offline following image analysis, or on the mobile phone using basic mage analysis software. The accuracy of the field method was limited by variability in image saturation, and variation in the illumination spectrum (lighting) and camera response. 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 in the field more widely, and that modest improvements in the consistency of photographic image collection and processing could yield measurements that are both accurate and precise.

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(Ahsan et al., 2006a; Ahsan et al., 2006b; Argos et al., 2010; Smith et al., 2000; Wasserman et al., 2004). 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 (Flanagan et al., 2012). 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(Van Geen et al., 2002). 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 (Pathey, 2009). 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, but don't provide continuous or discrete As concentrations, and often are subject to bias (George et al., 2012; Steinmaus et al., 2006; Van Geen et al., 2005).

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 (Carro Perez and Francisca, 2013; Jia et al., 2015; Kearns and Tyson, 2012; Paciornik et al., 2006; Salman et al., 2012). 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 into quantitative data, and apply this method in the field to measure As levels in rural Bangladesh, 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. It also identifies a number of sources of

3

uncertainty in color measurements that affect both the accuracy and the detection limits of colorimetric analysis.

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 (Ahsan et al., 2000; Argos et al., 2010; Dhar et al., 2011; Horneman et al., 2004). For this study, village health workers (VHW) sampled 376 tube wells that have been part of previous studies (Cheng et al., 2005; van Geen et al., 2014; Van Geen et al., 2002; Van Geen et al., 2007; Van Geen et al., 2003). 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. http://www.sensafe.com/) according to the manufactures protocol and previous studies (George et al., 2012; Rahman et al., 2002; Steinmaus et al., 2006; Van Geen et al., 2005) 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, http://www.dgkcolortools.com), (Figure SI-3 to Figure SI-8).

Digital image processing was conducted using Adobe Photoshop CS6 (Version 13.0 x64, 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. The Levels tool was used to carry out black normalization for the photo, by defining the black spot on the color chart as black (0,0,0) in RGB space. 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, for example for the color sensitive portion of the test kit strip. The average RGB color value of the pixels in the center of the test strip, an undeveloped test strip, and the standards of the EQ standard concentration chart were extracted and recorded for regressions of color values and As concentrations.

Once successful image processing routines were developed in Photoshop, we performed parallel color extraction on photographs of test strips taken in Bangladesh during testing within the phone directly. This inline photographic analyses were performed on an Android phone with the ColorMeter app v3.1.0 (http://vistechprojects.blogspot.com/). Readings in the Blue and Green channel were recorded at two spots on each test strip: (a) the center of the color-sensitive pad at the end of strip, and (b) a white section of the strip next to the pad. As a simplified normalizing

procedure, each reading for the reacting pad was multiplied by the ratio of 200, an arbitrary value comparable to the average, to the reading in the same channel for a nearby white section of the strip. Although this analysis was conducted after the field work, it could be applied in real-time in the field by calling the ColorMeter app (full version) from within a surveying app such as SurveyCTO (surveyCTO.com), a commercial adaptation of the open-source app ODK Collect (https://opendatakit.org/use/collect/). A SurveyCTO form in MS Excel that prompts the user to take a photograph of the test strip, calls on ColorMeter to obtain at two spots on the tests strip, and converts these readings to an estimated As concentration on the basis of an internal calibration is provided with the Supplementary Material.

Full spectrum visible-light diffuse reflectance data of the color of test strips under more controlled conditions using a spectrometer help to establish the theoretical basis for the measurements, and to identify potential limits to its application. Diffuse reflectance of the test strips of tap water spiked with known concentrations of As was measured using a Konica Minolta CR-700d Spectrophotometer (Konica Minolta Sensing Americas, Inc.) within the visible spectrum ranging from 360 to 750 nm. Diffuse reflectance (and absorbance) was recorded at 700, 550 and 400 nm, the centers of the R, G and B channels obtained in photographs to facilitate comparisons between photographic data and reflectance. 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 (Van Geen et al., 2007). 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 using a Thermo Fisher Scientific, Model: Element XR (Cheng et al., 2004).

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 center the data. A multiple linear regression between the normalized RGB color values and the known As concentrations was performed.

Results and Discussion

Standard Curve Development. There is a clear relationship between the visible diffuse reflectance spectrum of field kit test strips and As concentration (Figure 1, Supporting Information Figure SI-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. The large dynamic range of the blue channel at low levels makes it the most important determinant of As concentration in the linear regression model at low concentrations.

The variation in reflectance are exploited to develop standard curves for As concentrations based on multiple linear regression. The nonlinear behavior observed in color levels with concentration at higher concentrations implies that nonlinear or piecewise regressions could be used to allow for accurate quantitation over a greater linear range, but that the errors in large concentrations may also be higher due to scatter. At high concentrations, red and green channels are most important to accurately quantify for effective concentration measurement because they change more gradually over this range (with changes in particle size of the HgAs solid that is responsible for color development in the field test kit). Some of the scatter in these measurements can be eliminated by using ratios of colors (essentially transforming colors into L*a*b color space), in part because ratios are more stable to variation in color intensity caused by variable illumination or color response.

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).

A single standard curve for all laboratory measurements was surprisingly effective at predicting concentrations (r^2 = 0.90), and a single standard curve is preferable in that it allows a regression to be applied to field measurements without collecting standard curves in the field for each camera and lighting condition. Over the entire range, the individual standard curves had r^2 values >0.95, largely because they were the product of uniform image collection conditions, and a single camera. 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 limit for the multiple linear regression analysis method 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.

Field Measurement. 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 (Chang and Reid, 1996).

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 3) (George et al., 2012). 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 (Lin and Shum, 2001; Prats-Montalbán et al., 2011). 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.

Mobile Phone Processing. The ultimate goal of a mobile phone method is to be able to not only collect images, but also to analyze them in the field, and return results immediately to the user. To do so, strip images were analyzed using the Android app ColorMeter for a total of 217 wells. Raw readings for the white portion of the strip average 170±25 (1-sigma) and range from 120 to 255 for both Blue and Green. Only one reading in the Blue channel and three readings in Green channel reach saturation at 255. Raw readings for the reacting pad average 110+50 (range 0-214) in the Blue channel. After normalizing to a value of 200 for the white portion of the strip, scaled Blue readings average 130±50 and range from 0 to 200 (Figure 3). The scatter in the relationship between Blue readings and As concentrations is on the order of ±20 units for concentrations up to about 100 µg/L but increases markedly at higher concentrations. Adjusted Green readings average 180±20 and range from 90 to 210 (Figure SI-9). The scatter in the relationship between adjusted ColorMeter readings does not increase with As concentrations to the same degree in the Green channel as it does in the Blue channel, but the magnitude of the decline in Green readings is considerably smaller. These changes in Blue and Green channel intensity were able to predict As concentrations well, similar to Photoshop normalizations, although there is more scatter in the Colormeter calculated concentrations at low levels (Figure 3).

Public Health Implications. 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 4), although care should be used when using kriged data for public health reasons because the averaged data hides considerable local variance. This increased accuracy has two considerable advantages over existing methods: (1) mapping is with more accurate chemical data, for example using ICP-MS, is usually limited in spatial coverage and thus is not practical to accurately assess spatial variability, (2) low resolution field kits are easier to measure in large numbers but are sufficiently qualitative that they are useful for binning into safe and unsafe categories but are difficult to leverage into estimates of average or median As concentrations. Using this method, it is possible to accurately predict area concentrations, and to observe the transitions in concentration, or the changes in concentration that occur over time. As such, we expect improved versions of this field method would be able to measure concentrations over time even in a single well.

Results from this study highlight the potential for mobile phone photometry to accurately determine As concentrations in the field. Field-based image analysis was successfully performed, and it is interesting to consider the implications of doing so for the purposes of informing drinking water quality measurements. An important public health need is to properly categorize water as safe (less than a WHO guideline of 10 or Bangladesh guideline of 50 µg/L). This remains a challenge for this method in the field as performed, though laboratory measurements suggest that it is within reach with improved analysis methods. Visual inspection properly categorizes water when below 10 µg/L (34/34 samples below 10 µg/L were accurately identified, two with values of 11 and 17 µg/L As were also classified in the same group), and is nearly as effective to delineate As concentrations of 25 and 50 µg/L As. Similar levels of misclassification across the 10 and 50 µg/L thresholds have been reported previously (George et al. 2012; van Geen et al., 2014). For reasons that are unclear, the bias towards overestimating As concentrations below but close to 50 µg/L is larger in the present study. From a public health perspective, this bias serves to benefit public health because it provides a conservative estimate of As concentrations, and does not lead to unexpected exposures.

The continuous data obtained through image processing as performed does not improve the use of this method to categorize water based on 10 µg/L public health thresholds, but shows some benefit for the 50 µg/L threshold. The data indicate that visual readings are comparable for the Photoshop method but more reliable than ColorMeter at relatively low As levels. For the 32 visual kit readings of 0 and 10 µg/L that were consistent with laboratory measurements, calculated As exceed 10 µg/L for 5 samples. Similarly, concentrations of 8 samples calculated from the linear regression are below 10 µg/L within the set of 37 samples for which visual kit readings of 25 and 50 µg/L were consistent with the laboratory measurements. Visual color readings are therefore more reliable than ColorMeter readings for classifying wells relative to the 10 µg/L threshold. The 50 µg/L threshold is more challenging for visual categorization because the visual readings tend to systematically overestimate the As concentration at these levels. Above 50 µg/L, the ColorMeter readings misclassify 23 out 154 wells, in 8 cases by overestimating and in 15 cases by underestimating the As concentration. This is a smaller number than the 39 misclassifications based on visual kit readings. Visual readings should continue to be relied upon for classifying wells in the field until a better, possibly automated procedure to analyze the photograph of a test strip has been found.

This method can be improved by improving the consistency of image collection conditions, potentially improving the utility of the method for continuous analysis and potentially improving the accuracy and precision of analysis. These methods are also not limited to As, and could be applied to many other colorimetric kits based on either colored solid substrate and aqueous solutions. In the future, the method should improve as phones with higher resolution and dynamic range are introduced. We have demonstrated that this method can be applied successfully within a mobile phone-based data collection environment to automate analysis and allow automated data aggregation and presentation, and to allow for real-time feedback with the end users, many of whom are directly impacted by these measurements. Improved automation of our method within a single mobile phone app could further enable widespread high quality data acquisition by individual stakeholders and citizen scientists.

Acknowledgement

References
Ahsan H, Chen Y, Parvez F, Argos M, Hussain AI, Momotaj H, et al. Health Effects of Arsenic Longitudinal Study (HEALS): description of a multidisciplinary epidemiologic investigation. Journal of Exposure Science and Environmental Epidemiology 2006a; 16: 191-205.

Ahsan H, Chen Y, Parvez F, Zablotska L, Argos M, Hussain I, et al. Arsenic exposure from drinking water and risk of premalignant skin lesions in Bangladesh: baseline results from the Health Effects of Arsenic Longitudinal Study. American Journal of Epidemiology 2006b; 163: 1138-1148.

Ahsan H, Perrin M, Rahman A, Parvez F, Stute M, Zheng Y, et al. Associations between drinking water and urinary arsenic levels and skin lesions in Bangladesh. Journal of Occupational and Environmental Medicine 2000; 42: 1195-1201.

Argos M, Kalra T, Rathouz PJ, Chen Y, Pierce B, Parvez F, et al. Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): a prospective cohort study. The Lancet 2010; 376: 252-258.

Carro Perez ME, Francisca FM. Digital analysis technique for uncertainty reduction in colorimetric arsenic detection method. Journal of Environmental Science and Health, Part A 2013; 48: 191-196.

Chang Y-C, Reid JF. RGB calibration for color image analysis in machine vision. IEEE Transactions on image processing 1996; 5: 1414-1422.

Cheng Z, Van Geen A, Seddique A, Ahmed K. Limited temporal variability of arsenic concentrations in 20 wells monitored for 3 years in Araihazar, Bangladesh. Environmental Science & Technology 2005; 39: 4759-4766.

Cheng Z, Zheng Y, Mortlock R, Van Geen A. Rapid multi-element analysis of groundwater by high-resolution inductively coupled plasma mass spectrometry. Analytical and bioanalytical chemistry 2004; 379: 512-518.

Dhar RK, Zheng Y, Saltikov CW, Radloff KA, Mailloux BJ, Ahmed KM, et al. Microbes enhance mobility of arsenic in Pleistocene aquifer sand from Bangladesh. Environmental science & technology 2011; 45: 2648-2654.

Flanagan SV, Johnston RB, Zheng Y. Arsenic in tube well water in Bangladesh: health and economic impacts and implications for arsenic mitigation. Bulletin of the World Health Organization 2012; 90: 839-846.

George CM, Zheng Y, Graziano JH, Rasul SB, Hossain Z, Mey JL, et al. Evaluation of an arsenic test kit for rapid well screening in Bangladesh. Environmental science & technology 2012; 46: 11213-11219.

Horneman A, van Geen A, Kent DV, Mathe P, Zheng Y, Dhar R, et al. Decoupling of As and Fe release to Bangladesh groundwater under reducing conditions. Part I: Evidence from sediment profiles. Geochimica et Cosmochimica Acta 2004; 68: 3459-3473.

Jia M-Y, Wu Q-S, Li H, Zhang Y, Guan Y-F, Feng L. The calibration of cellphone camera-based colorimetric sensor array and its application in the determination of glucose in urine. Biosensors and Bioelectronics 2015; 74: 1029-1037.

Kearns J, Tyson J. Improving the accuracy and precision of an arsenic field test kit: increased reaction time and digital image analysis. Analytical Methods 2012; 4: 1693-1698.

Lin S, Shum H-Y. Separation of diffuse and specular reflection in color images. Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. 1. IEEE, 2001, pp. I-341-I-346 vol. 1.

Paciornik S, Yallouz AV, Campos RC, Gannerman D. Scanner image analysis in the quantification of mercury using spot-tests. Journal of the Brazilian Chemical Society 2006; 17: 156-161.

Pathey P. Monitoring the Situation of Children and Women: Multiple Indicator Cluster Survey 2009. Bangladesh Bureau of Statistics and United Nations Children's Fund (UNICEF) 2009; 1.

Prats-Montalbán JM, De Juan A, Ferrer A. Multivariate image analysis: a review with applications. Chemometrics and Intelligent Laboratory Systems 2011; 107: 1-23.

Rahman MM, Mukherjee D, Sengupta MK, Chowdhury UK, Lodh D, Chanda CR, et al. Effectiveness and reliability of arsenic field testing kits: are the million dollar screening projects effective or not? Environmental science & technology 2002; 36: 5385-5394.

Salman M, Athar M, Shafique U, Anwar J, Rehman R, Ameer S, et al. Micro-determination of arsenic in aqueous samples by image scanning and computational quantification. Analytical Methods 2012; 4: 242-246.

Smith AH, Lingas EO, Rahman M. Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. Bulletin of the World Health Organization 2000; 78: 1093-1103.

Steinmaus CM, George CM, Kalman DA, Smith AH. Evaluation of two new arsenic field test kits capable of detecting arsenic water concentrations close to 10 μg/L. Environmental science & technology 2006; 40: 3362-3366.

van Geen A, Ahmed EB, Pitcher L, Mey JL, Ahsan H, Graziano JH, et al. Comparison of two blanket surveys of arsenic in tubewells conducted 12years apart in a 25km 2 area of Bangladesh. Science of the Total Environment 2014; 488: 484-492.

Van Geen A, Ahsan H, Horneman AH, Dhar RK, Zheng Y, Hussain I, et al. Promotion of well-switching to mitigate the current arsenic crisis in Bangladesh. Bulletin of the World Health Organization 2002; 80: 732-737.

Van Geen A, Cheng Z, Jia Q, Seddique AA, Rahman MW, Rahman MM, et al. Monitoring 51 community wells in Araihazar, Bangladesh, for up to 5 years: Implications for arsenic mitigation. Journal of Environmental Science and Health Part A 2007; 42: 1729-1740.

Van Geen A, Cheng Z, Seddique A, Hoque M, Gelman A, Graziano J, et al. Reliability of a commercial kit to test groundwater for arsenic in Bangladesh. Environmental science & technology 2005; 39: 299-303.

Van Geen A, Zheng Y, Versteeg R, Stute M, Horneman A, Dhar R, et al. Spatial variability of arsenic in 6000 tube wells in a 25 km2 area of Bangladesh. Water Resources Research 2003; 39.

Wasserman GA, Liu X, Parvez F, Ahsan H, Factor-Litvak P, van Geen A, et al. Water arsenic exposure and children's intellectual function in Araihazar, Bangladesh. Environmental health perspectives 2004: 1329-1333.

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 µg/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 µg/L, and full range. Field blanks were classified by ICP-MS determined concentrations below 1 µg/L. If there were less than 5 blanks, the detection limit was not provided.

| Date | Label | Figure | 0-150 µg/L | | | Full range | | | |
			n	r^2	DL (µg/L)	n	r^2	DL (µg/L)	Max As (ug/L)
April 16th, 2015	Lab 1	SI-1	5	1.00	-	9	0.96	-	1000
April 23rd, 2015	Lab 2	SI-2	13	0.96	-	17	0.99	-	1000
Oct, 7th, 2015	Lab 3	SI-3	14	0.97	-	24	0.95	-	1280
Jul 30th, 2016	Lab 4	SI-4	10	0.99	8.2	13	0.97	150	1000
2015-2016	Lab 1-4	SI-5	42	0.96	9.2	62	0.90	180	1280
July, 2015	Field	SI-6	220	0.83	22	274	0.75	61	750

Figure 1: Relationshiip 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 Correected, and C) B Corrected.
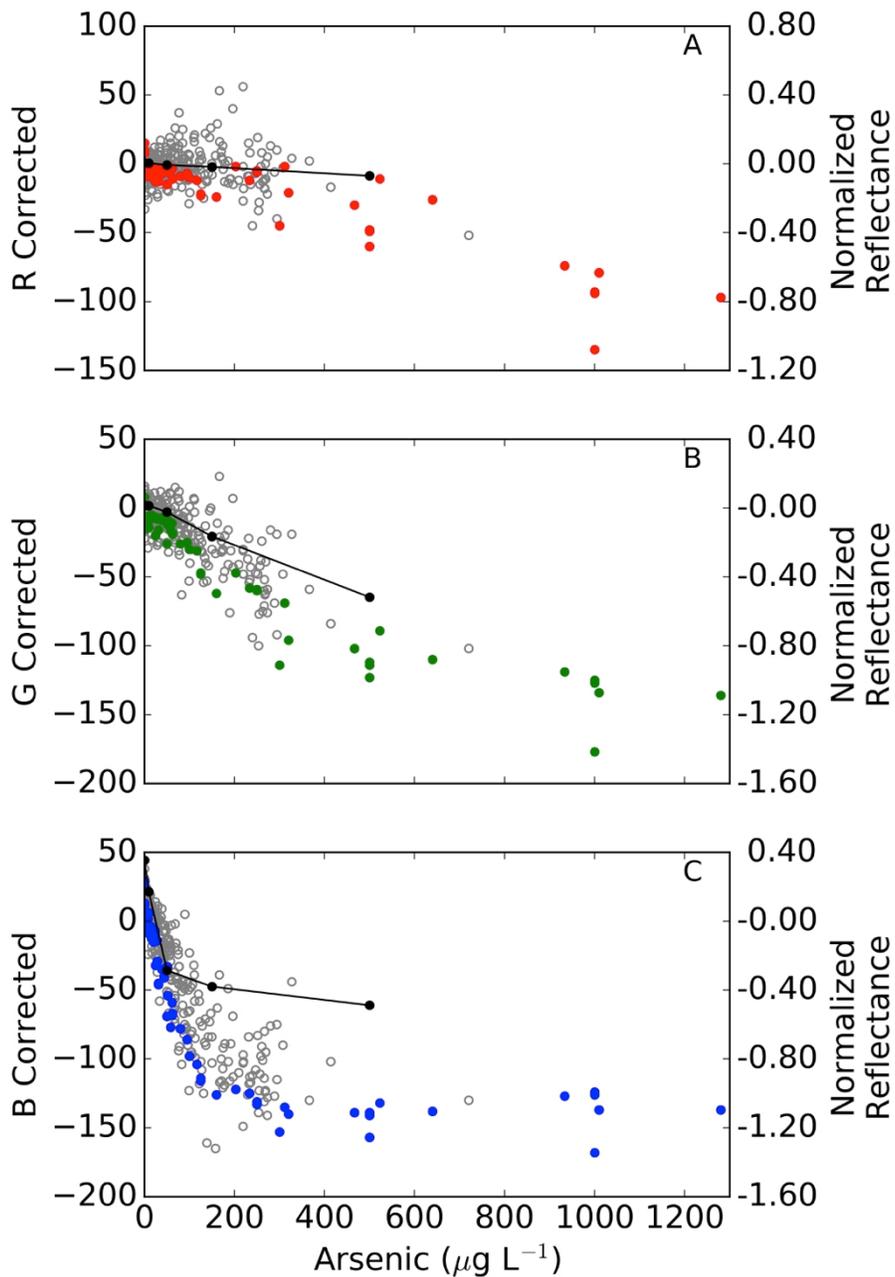
Figure 2: Cross validation plot of photo-predicted As concentration compared with lab-measured As concentrations in laboratory samples (red) and field samples (blue) using a single, combined multiple regression model of Photoshop-extracted RGB color data.
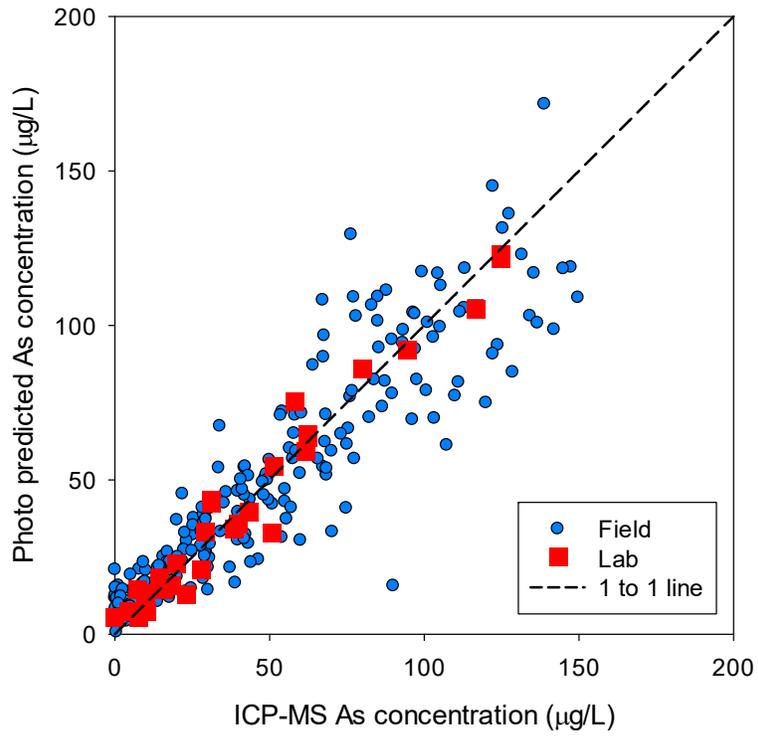
Figure 3: Cross validation plots showing the comparison of field measured categorical As concentrations (left), Photoshop predicted field concentrations (center) and ColorMeter-predicted concentrations (right) compared to As concentrations measured by ICP-MS.
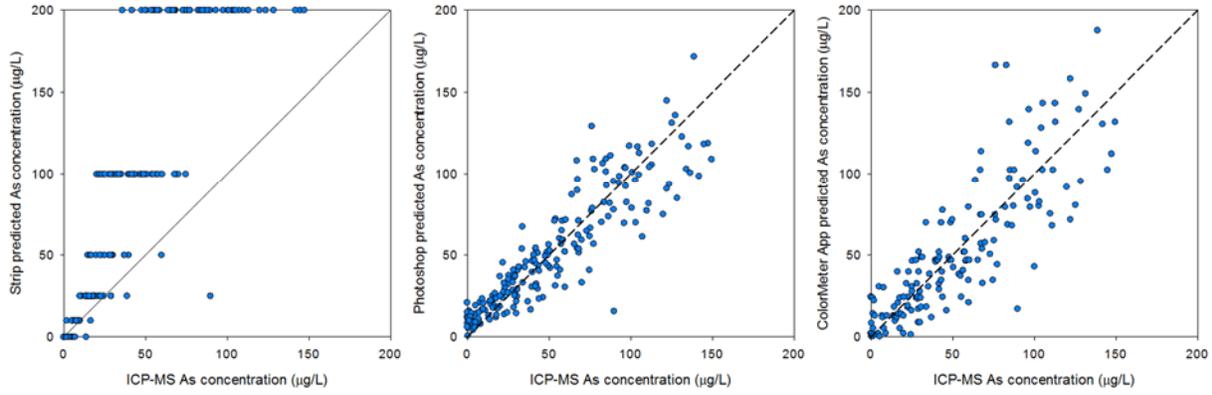
Figure 4: Contour plots of a) ICPMS determined b) field kit photo predicted and c) field kit strip predicted As concentration expressed in μg/L. The location and As concentrations of wells sampled are displayed in d. Well with depths >50 m are not shown.
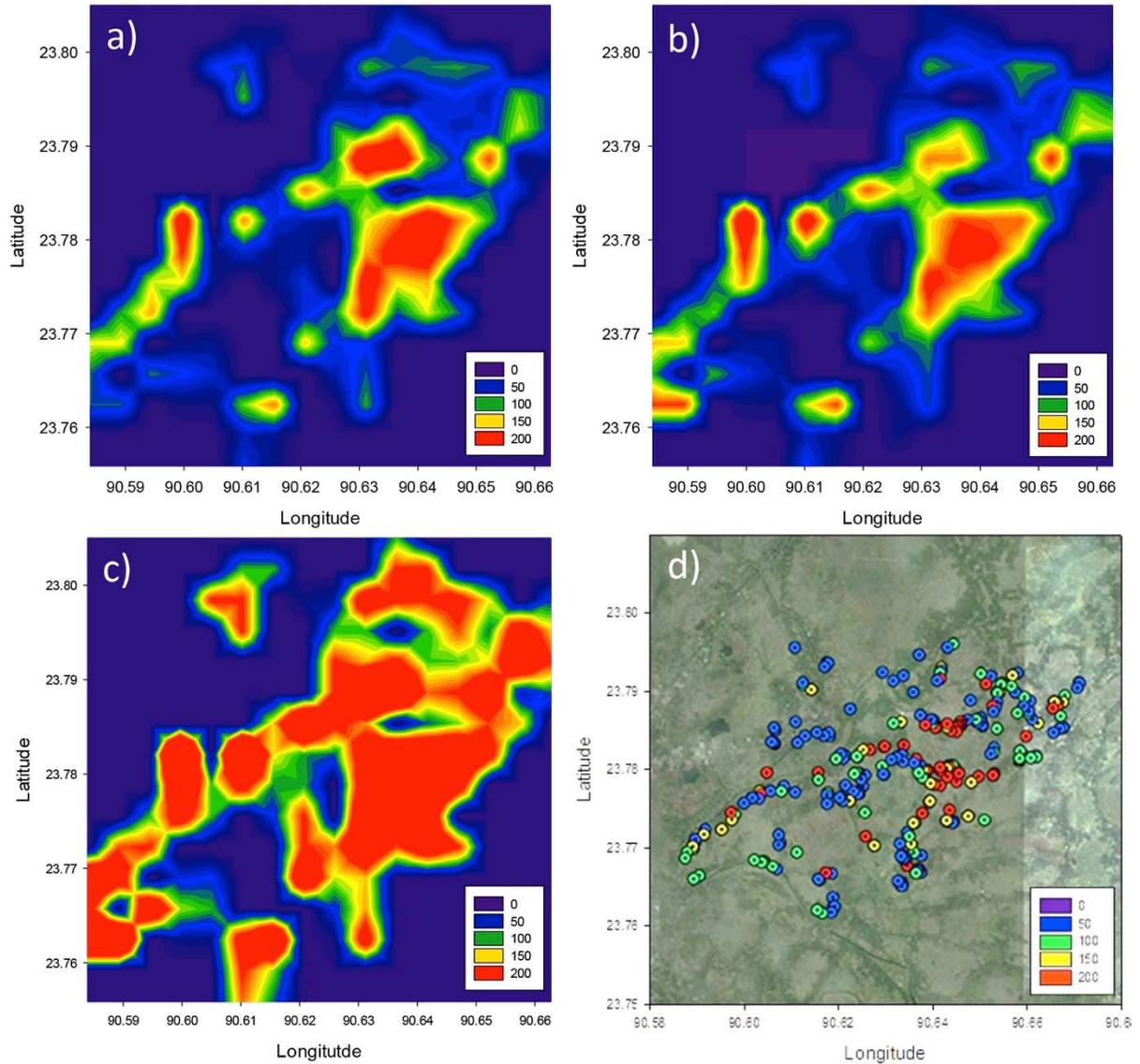
1 **Supporting Information**
2
3
4 # Quantitative Drinking Water Arsenic Concentrations in Field
5 # Environments Using Mobile Phone Photometry of Field Kits
6

7 EZAZUL HAQUE[a,b], BRIAN J. MAILLOUX[b], DAISY DE WOLFF[b], SABINA GILIOLI[b],
8 COLETTE KELLY[b], ERSHAD AHMED [c], CHRISTOPHER SMALL[a], KAZI MATIN
9 AHMED[c], ALEXANDER VAN GEEN[a], BENJAMIN C. BOSTICK[a]
10
11 [a] *Lamont-Doherty Earth Observatory of Columbia University, Route 9 W, Palisades, New York*
12 *10964, United States*
13 [b] *Department of Environmental Sciences, Barnard College, 3009 Broadway, New York, New*
14 *York 10027, United States*
15 [c] *Department of Geology, Dhaka University, Dhaka 1000, Bangladesh*
16
17

18   Figure SI-1: Absorbance spectrum of field kit test strips of a blank and tests used for lab
19   solutions of As concentrations of 10, 50, 150, and 500 µg/L.
20
21

22
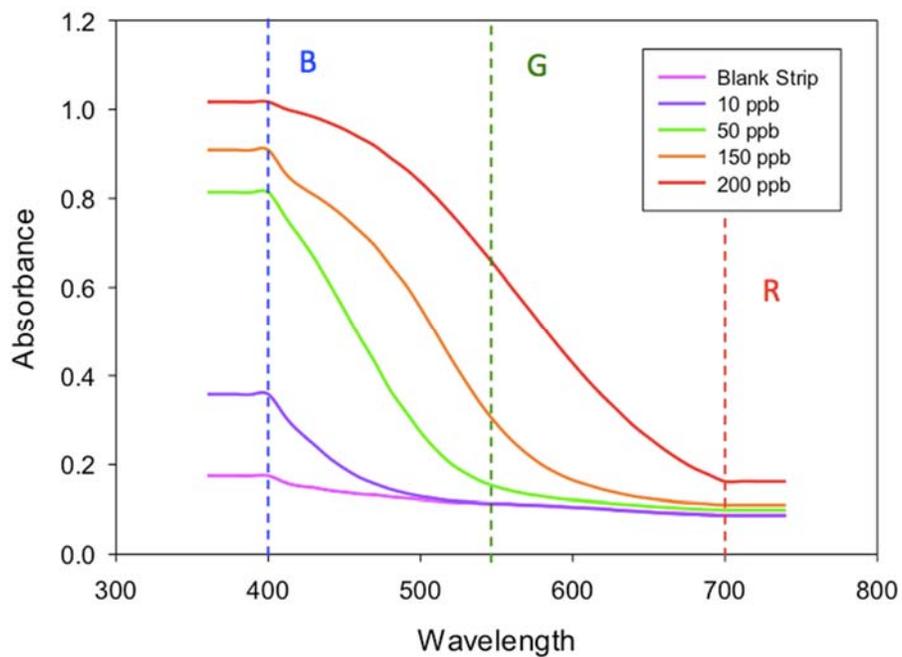
Figure SI-2: Relationship between ICP-MS measured As concentration (µg/L) for RGB color
values a) not normalized and b) normalized. Field and lab samples are represented by gray and
blue points respectively.
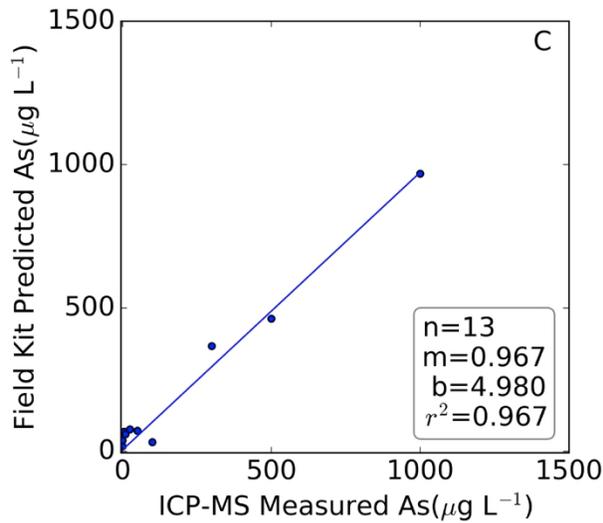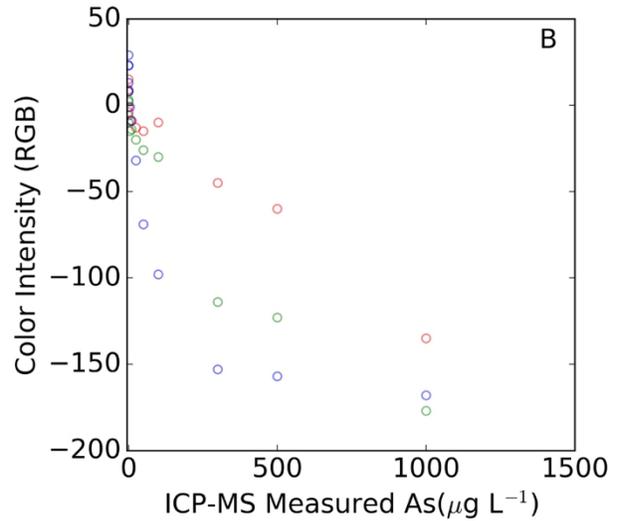
a)   b)

3
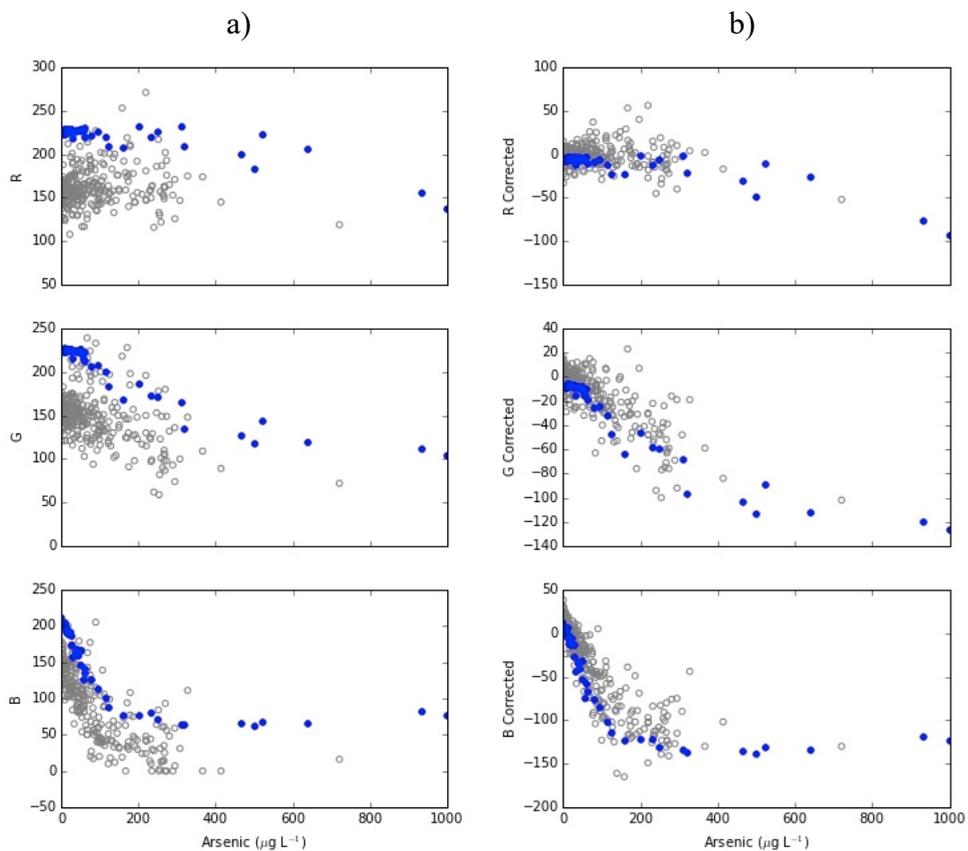
32 Figure SI-3: Summary of experiment done on April 16th, 2015 in laboratory controlled
33 conditions showing the black normalized photograph taken on an iPhone 5S (A),
34 relationship between normalized R, G, B with respect to As concentration (B), inductively
35 coupled plasma mass spectrometry (ICP-MS) determined versus photo predicted As
36 concentration for full range up to 1500 µg/L, and an expanded version of (C) is shown in (D)
37 for As concentrations up to 150 µg/L.
38
39
40
41
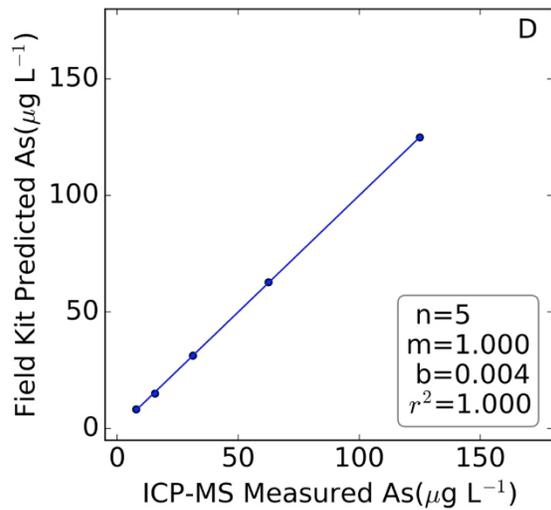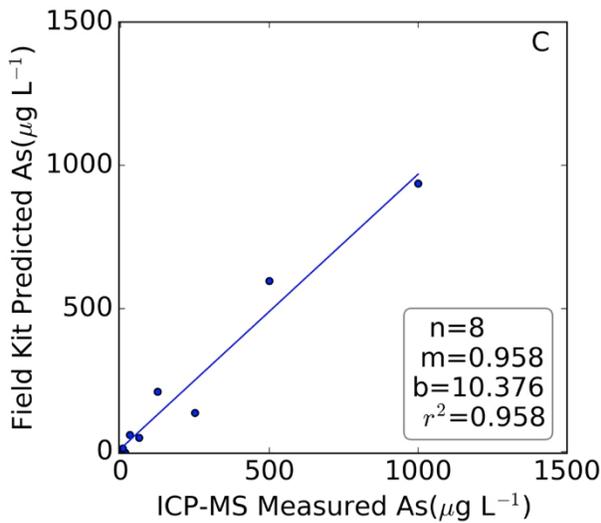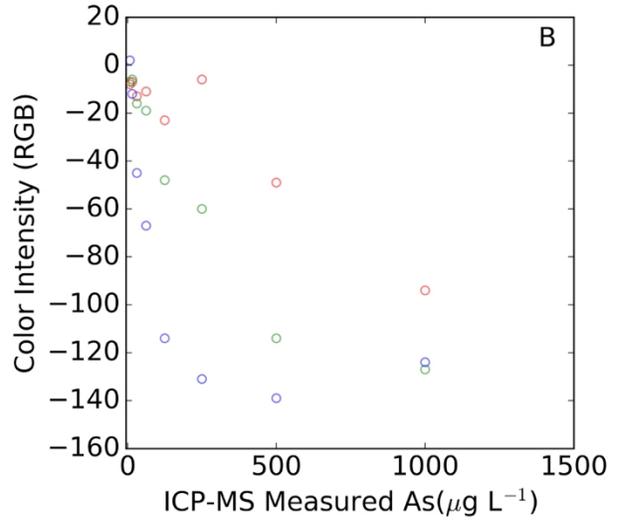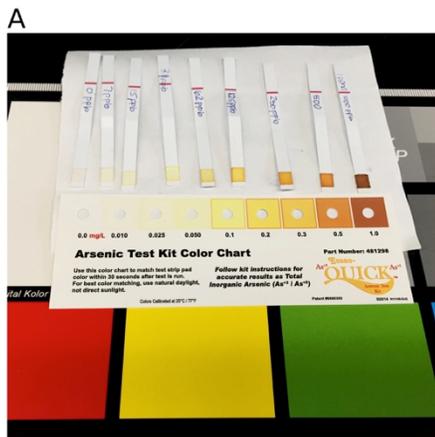
42
43

44   Figure SI-4: Summary of experiment done on April 23rd, 2015 in laboratory controlled
45   conditions showing the black normalized photograph taken on an iPhone 5S (A),
46   relationship between normalized R, G, B with respect to As concentration (B), inductively
47   coupled plasma mass spectrometry (ICP-MS) determined versus photo predicted As
48   concentration for full range up to 1500 μg/L, and an expanded version of (C) is shown in (D)
49   for As concentrations up to 150 μg/L.
50
51
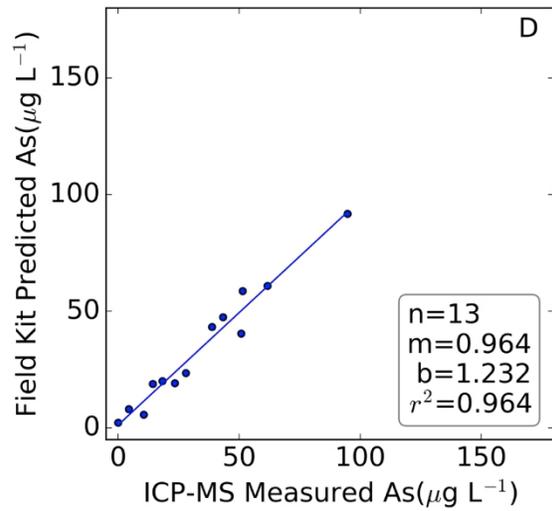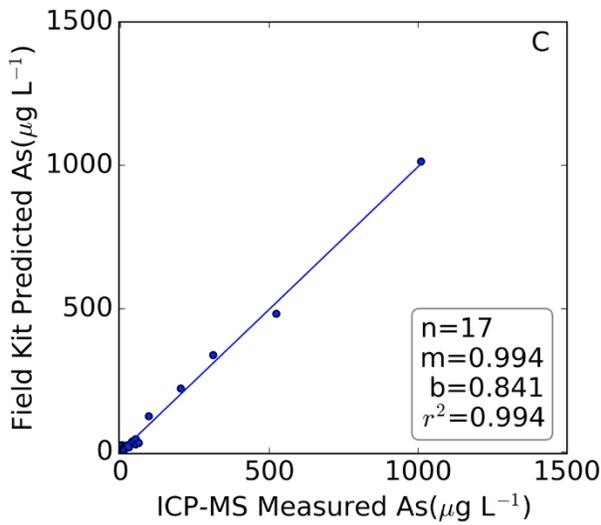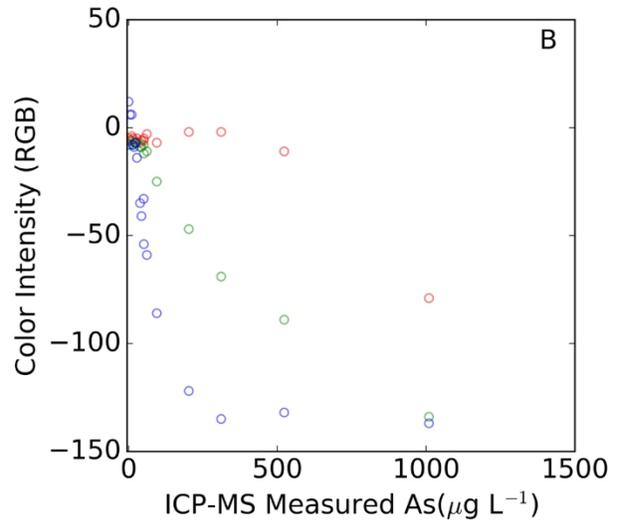
52

53 Figure SI-5: Summary of experiment done on October 7th, 2015 in laboratory controlled
54 conditions showing the black normalized photograph taken on an iPhone 5S (A), relationship
55 between normalized R, G, B with respect to As concentration (B), inductively coupled plasma
56 mass spectrometry (ICP-MS) determined versus photo predicted As concentration for full range
57 up to 1500 µg/L, and an expanded version of (C) is shown in (D) for As concentrations up to 150
58 µg/L.
59
60
61
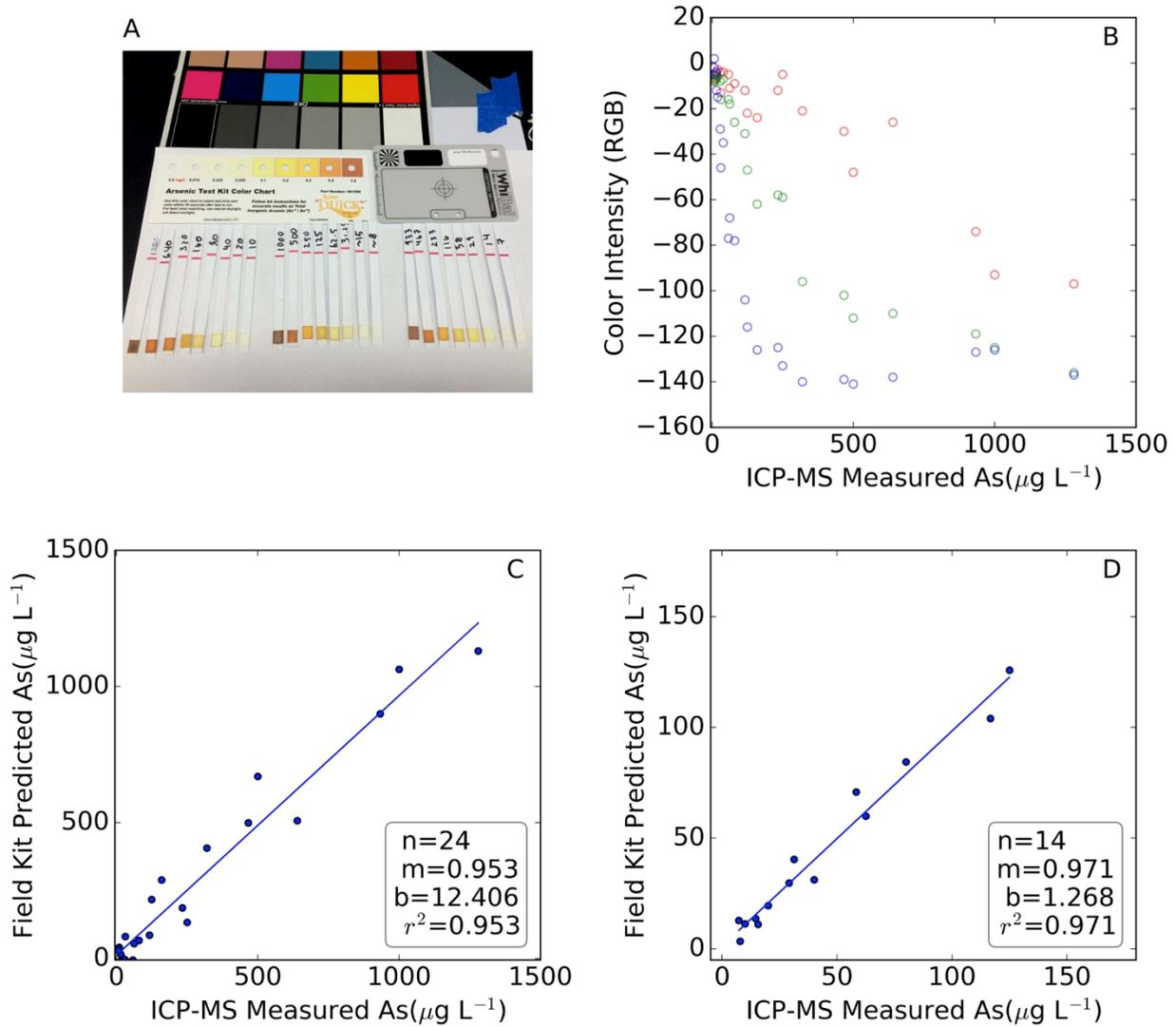
62
63

64 Figure SI-6: Summary of experiment done on October 7th, 2015 in laboratory controlled
65 conditions showing the black normalized photograph taken on an iPhone 5S (A), relationship
66 between normalized R, G, B with respect to As concentration (B), inductively coupled plasma
67 mass spectrometry (ICP-MS) determined versus photo predicted As concentration for full range
68 up to 1500 µg/L, and an expanded version of (C) is shown in (D) for As concentrations up to 150
69 µg/L.
70
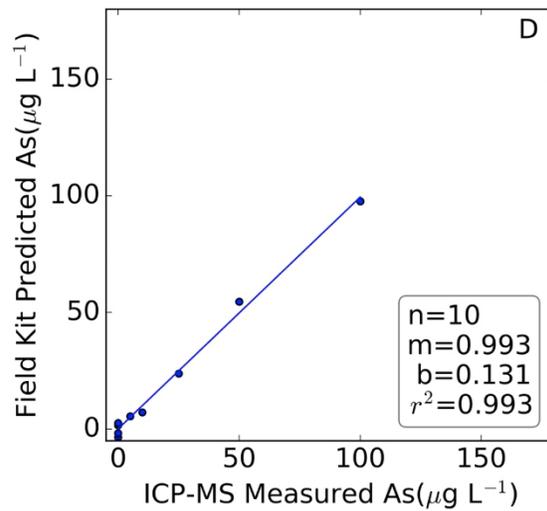
71
72
73
74

75 Figure SI-7: Summary of experiment done in laboratory controlled conditions from Figure
76 SX-SX showing the black normalized photograph taken on an iPhone 5S (A), relationship
77 between normalized R, G, B with respect to As concentration (B), inductively coupled
78 plasma mass spectrometry (ICP-MS) determined versus photo predicted As concentration
79 for full range up to 1500 μg/L, and an expanded version of (C) is shown in (D) for As
80 concentrations up to 150 μg/L.
81
82

83
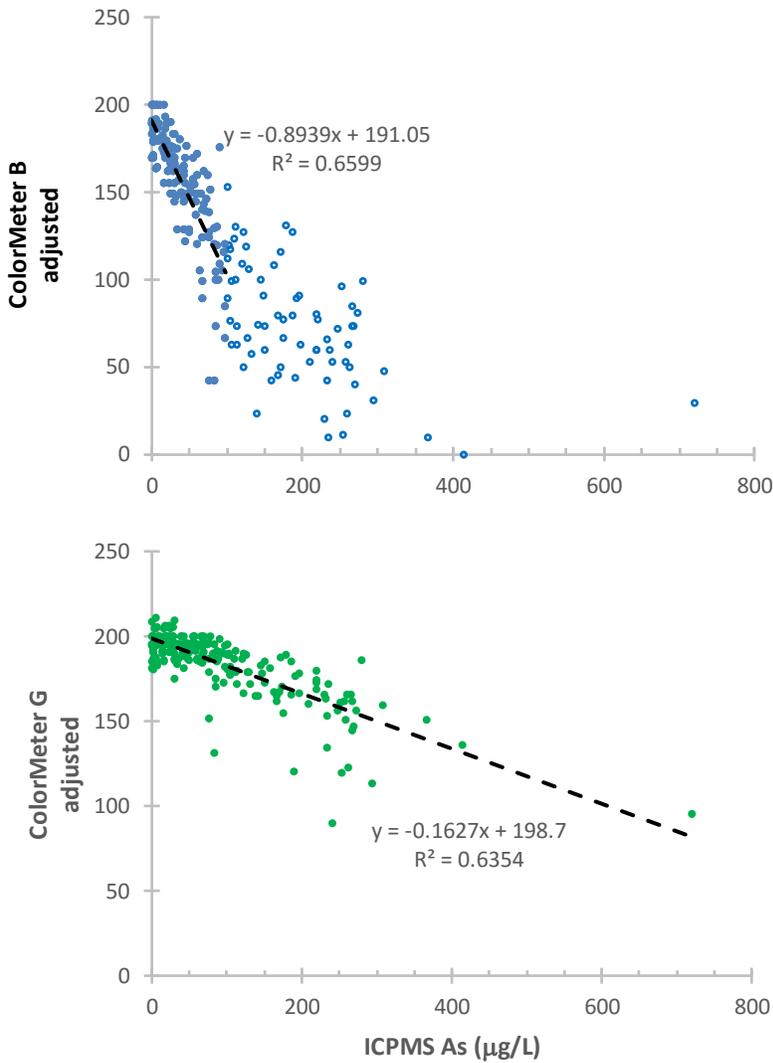84
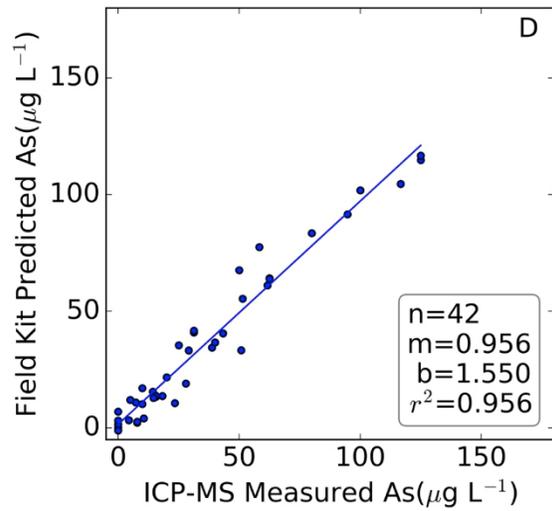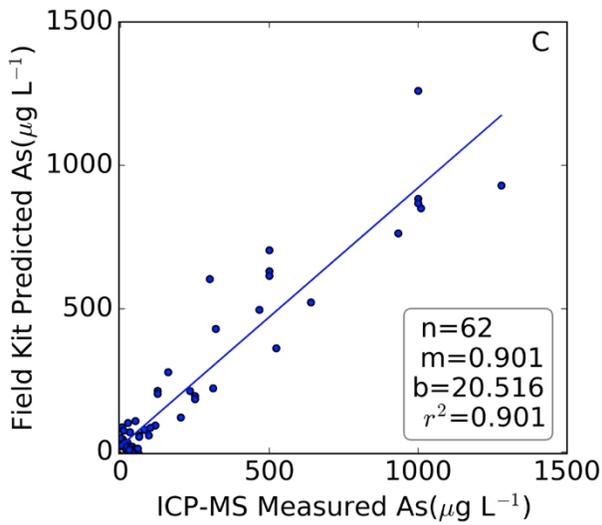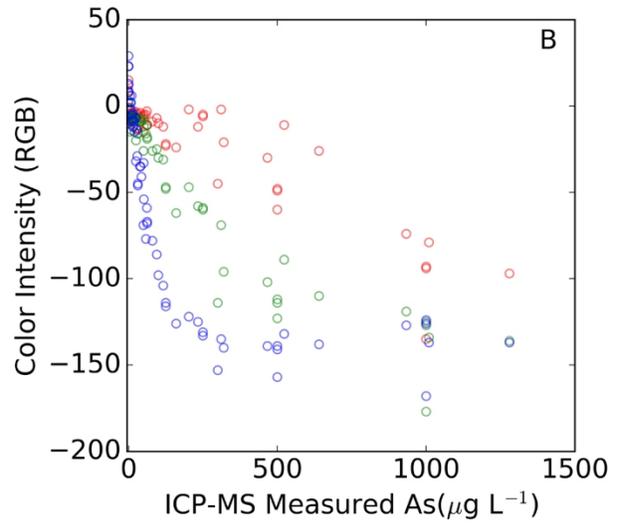
Figure SI-8: Summary of tests done in field in Araihazar, Bangladesh on July-2015 showing the black normalized photograph taken on an iPhone 5S (A), relationship between normalized R, G, B with respect to As concentration (B), inductively coupled plasma mass spectrometry (ICP-MS) determined versus photo 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.
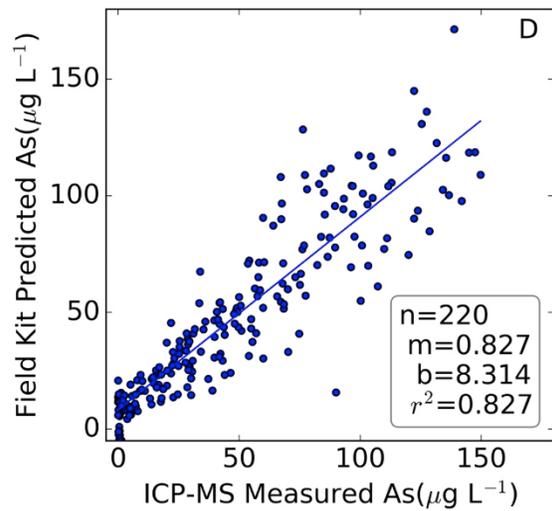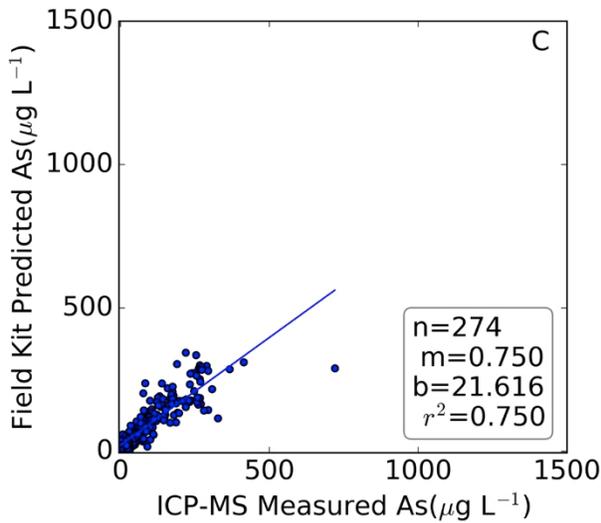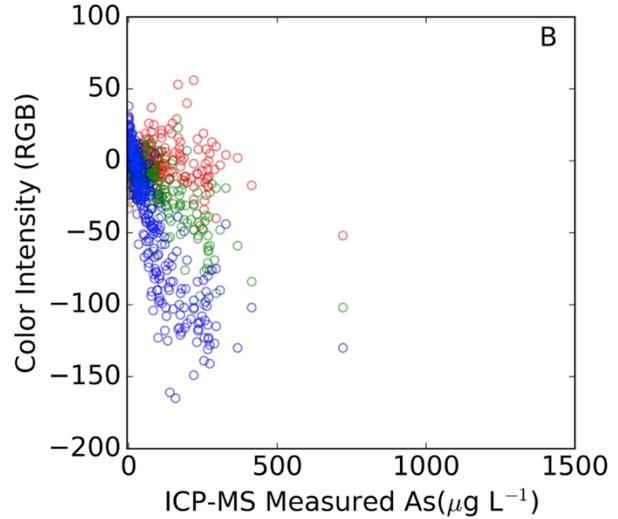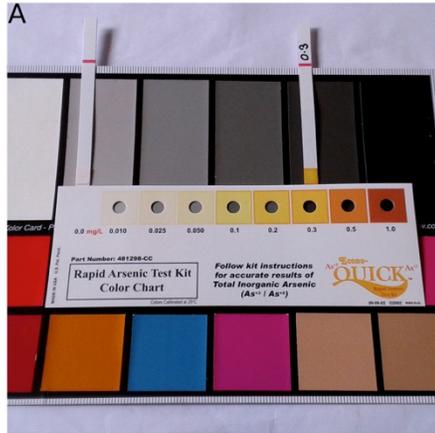
95
96 Figure SI-9. ColorMeter readings for (a) Red and (b) Green as a function of As
97 concentrations in 217 well-water samples from Bangladesh first tested in the field with the
98 kit and subsequently analyzed in the laboratory by ICPMS. Readings for the reacting
99 portion of the strip that was photographed in the field immediately after testing are
100 adjusted by the proportion required to yield a constant R and G readings of 200 for a white
101 portion of the test strip. The regression line shown in (a) considers only samples
102 containing up to 100 μg/L As.
103
104

105
106
107