Reliability of a Commercial Kit To Test Groundwater for Arsenic in Bangladesh

A. VAN GEEN, *1 Z. CHENG, †
A. A. SEDDIQUE, ‡ M. A. HOQUE, †
A. GELMAN, § J. H. GRAZIANO, ‡
H. AHSAN, § F. PARVEZ, ‡ AND
K. M. AHMED †
Lamont-Doherty Earth Observatory of Columbia University,
Palisades, New York 10964, Department of Geology, University
of Dhaka, Bangladesh, Department of Statistics, Columbia
University, New York, and Mailman School of Public Health,
Columbia University, New York

A comparison of field and laboratory measurements of
arsenic in groundwater of Araihazar, Bangladesh, indicates
that the most widely used field kit correctly determined
the status of 88% of 799 wells relative to the local standard
of 50 μg/L As. Additional tests show that the inconsistencies,
mainly underestimates in the 50–100 μg/L As range, can
be avoided by increasing the reaction time from 20 to
40 min. Despite this limitation, the field data already
compiled for millions of wells by the Bangladesh Arsenic
Mitigation and Water Supply Project, in combination with
information on well location and depth, should prove to
be extremely useful to prioritize interventions in thousands
of affected villages.

Introduction

A compilation of recent estimates suggests that over 100
million villagers of Bangladesh, West Bengal (India), Vietnam,
China, and several other South Asian countries drink and
cook with groundwater drawn from shallow wells containing
over 10 μg/L As, the World Health Organization guideline
value for arsenic in drinking water (1–4). In addition, one-
third of the wells in Bangladesh producing water with less
than 10 μg/L As do not meet the WHO guideline for Mn in
drinking water of 50 μg/L (3). Technically plausible ap-
proaches to mitigate the situation have not been tested on
anything beyond the pilot scale (5). A viable strategy to provide
safe water throughout rural Bangladesh and other affected
countries has yet to be articulated, let alone implemented.
Even the value of past testing campaigns to identify safe
and unsafe wells with field kits, the logical first phase of any
mitigation effort, has been questioned by an influential group
because of purported discrepancies between field kit results
and laboratory measurements (6).

This report strikes a more optimistic note on the basis of
the experiences of an interdisciplinary team of health, social,
and earth scientists working since early 2000 in a 25 km²
study area that encompasses ~6600 tube wells and a
population of 70 000 in Araihazar Upazila, Bangladesh (~
8). We focus here on the performance of a field kit for As that
has been used to test millions of wells in Bangladesh since
its introduction by the Hach Company (Loveland, CO) in
2001. NGO workers contracted by the Bangladesh Arsenic
Mitigation and Water Supply Project (BAMWSP) used the
Hach kit to test all tube wells in our 25 km² study area in
2003. Groundwater from many of the same wells had been
sampled in 2000–2001 and analyzed in our laboratories by
groundwater. The generated arsine gas (AsH 3) is entrained
in acid-cleaned bottles after pumping for ~5 min and acidified
in the field to 1% HCl (Seastar, Fisher Scientific). The addition
reduced the pH to the point where Fe oxhydroxides do not
precipitate because the amount of acid in the sample (0.12
N) exceeds by a factor of 6 the highest alkalinitities reported
for Bangladesh groundwater (3). Steps taken to determine
the reliability of the sampling procedure and analyses by
GFAAS for the 2000–2001 samples have been described
elsewhere (8). A similar quality control procedure was
followed in 2003 with a subset of 71 wells where a replicate
sample was collected and a third sample bottle was spiked
in the field for recovery tests with a known amount of As.
Well water was analyzed for As and 30 other constituents by
high-resolution inductively coupled plasma mass spectrom-
etry (HR ICP-MS) following 1:10 dilution. The analytical
detection limit of this method for As is 0.1 μg/L. The minimum
concentration of dissolved As that is reliably determined
by the entire procedure is estimated to be 1 μg/L, taking into
account small quantities of As released upon acidification
by fine aquifer particles that are occasionally collected with
a sample. In addition to all 993 samples collected in 2003,
groundwater collected in 2000–2001 from the subset of 349
wells sampled twice was analyzed by HR ICP-MS.

Field Kit Measurements. BAMWSP-trained NGO workers
used the two-step Hach EZ arsenic kit (product 2822800) to
analyze groundwater from all wells in the upazila in 2002–
2003. The method is one of several existing variants of the
1879 Gutzeit method (10) and involves the addition of
prepackaged sulfurous acid and zinc powder to ~50 mL of
groundwater. The generated arsine gas (AsH 3) is entrained
with H 2 bubbles emanating from the acidified sample and
trapped by a strip of paper impregnated with mercuric
bromide. The option to trap sulfide with a lead acetae-
impregnated cotton ball was taken out of the kit because
sulfide levels are generally too low in Bangladesh groundwater

Methods

Sample Collection for Laboratory Analysis. Groundwater
was collected and analyzed in the laboratory from 4999 wells
in March–June 2000, an additional 972 wells in November–
December 2001, and 993 wells in April 2003, of which 349
had already been sampled in 2000 or 2001. To re-identify
individual wells, geographic coordinates were determined
with handheld global positioning system (GPS) receivers and
numbered stainless steel plates were attached to the base of
each well pump on the first sampling occasion.

All well-water samples were collected without filtration
in acid-cleaned bottles after pumping for ~5 min and acidified
in the field to 1% HCl (Seastar, Fisher Scientific). The addition
reduced the pH to the point where Fe oxhydroxides do not
precipitate because the amount of acid in the sample (0.12
N) exceeds by a factor of 6 the highest alkalinitities reported
for Bangladesh groundwater (3). Steps taken to determine
the reliability of the sampling procedure and analyses by
GFAAS for the 2000–2001 samples have been described
elsewhere (8). A similar quality control procedure was
followed in 2003 with a subset of 71 wells where a replicate
sample was collected and a third sample bottle was spiked
in the field for recovery tests with a known amount of As.
Well water was analyzed for As and 30 other constituents by
high-resolution inductively coupled plasma mass spectrom-
etry (HR ICP-MS) following 1:10 dilution. The analytical
detection limit of this method for As is 0.1 μg/L. The minimum
concentration of dissolved As that is reliably determined
by the entire procedure is estimated to be 1 μg/L, taking into
account small quantities of As released upon acidification
by fine aquifer particles that are occasionally collected with
a sample. In addition to all 993 samples collected in 2003,
groundwater collected in 2000–2001 from the subset of 349
wells sampled twice was analyzed by HR ICP-MS.

Field Kit Measurements. BAMWSP-trained NGO workers
used the two-step Hach EZ arsenic kit (product 2822800) to
analyze groundwater from all wells in the upazila in 2002–
2003. The method is one of several existing variants of the
1879 Gutzeit method (10) and involves the addition of
prepackaged sulfurous acid and zinc powder to ~50 mL of
groundwater. The generated arsine gas (AsH 3) is entrained
with H 2 bubbles emanating from the acidified sample and
trapped by a strip of paper impregnated with mercuric
bromide. The option to trap sulfide with a lead acetae-
impregnated cotton ball was taken out of the kit because
sulfide levels are generally too low in Bangladesh groundwater

10.1021/es0491073 CCC: $30.25
© xxx American Chemical Society
Published on Web 08/30/2000
PAGE EST: 4.2
VOL. xx, NO. xx, xxxx / ENVIRON. SCI. & TECHNOL. • A
to cause significant interference. BAMWSP evidently judged, in our opinion correctly, that the very hypothetical sulfide interference was not worth the added complication and the risk of jeopardizing the results by wetting the test strip. After the 20-min reaction time stipulated by the manufacturer, the color of the orange-brown circle on the strip is compared visually to a reference scale showing readings corresponding to As concentrations of 0, 10, 25, 50, 100, 250, and 500 μg/L.

For 799 out of 993 wells sampled in April 2003, the color of the paint applied on the spout of each well could be determined directly or by asking well owners for their card with the test result. Few if any traces of the paint adhered to the rusty spouts by October 2003 however.

Our own testing with the Hach kit had suggested that the field measurement might be sensitive to how long the sample is allowed to react. For this reason, a subset of wells was retested with the kit by following the recommended procedure and by doubling the reaction time to 40 min. A first set of 43 wells was selected randomly in April and June 2003; another 65 wells were retested in the field September 2003 specifically because of observed discrepancies between the laboratory and BAMWSP results. The total of 108 wells that were retested by our team may therefore be biased toward the type of groundwater that is difficult to analyze with the Hach kit.

Results

Quality Control of Laboratory Measurements. For groundwater analyzed by HR ICP-MS containing <1–600 μg/L As, differences in concentration for 71 sets of replicates averaged 0 ± 5 μg/L (Figure 1). The difference in As concentration between the average of the two replicates and the corresponding spiked sample averaged 472 ± 22 μg/L, which corresponds to a recovery of 96 ± 4% relative to the expected value of 492 μg/L. Following the model derived in the Appendix of ref 8, the standard error of individual measurements by HR ICP-MS is estimated from the expression: $\sigma_{\text{meas}} = \sqrt{\sigma_{\text{cal}}^2 + \sigma_{\text{meas}}^2}$, where the As concentration is $\theta$, the single measurement error $\sigma_{\text{meas}} = (5 \mu g/L)/\sqrt{2} = 4 \mu g/L$ is based on the reproducibility of replicates, and $\sigma_{\text{cal}}$ is the relative calibration error of 0.02 derived from the reproducibility of the recovery tests corrected for variability in the dilution of the spike. The standard error for all HR ICP-MS measurements presented in this study was calculated according to this expression. An absolute uncertainty of 4 μg/L obtained from replicate analyses dominates below a concentration of 150 μg/L. Above this concentration, a relative error of ∼2% because of calibration errors or minor matrix effects becomes the larger source of uncertainty.

Comparison of As Concentrations in 2000–2001 and 2003. Relative to the Bangladesh standard of 50 μg/L for As in drinking water, the status of 331 (95%) out of 349 wells was unaffected by the outcome of the second sampling and analysis. Arsenic concentrations measured by HR ICP-MS were consistent for 318 (91%) of the 349 pairs of samples (Figure 2). Overlap of error bars corresponding to 3 times the standard error for individual HR ICP-MS measurements is the criterion that was used to determine consistency. The probability that As concentrations measured in the other 31 pairs of samples that do not meet this criterion are actually indistinguishable is therefore extremely low. The differences cannot be attributed to changes in As concentrations over time in the 2000–2001 sample bottles since previous determinations by GFAA are consistent with re-analysis by HR ICP-MS for all 31 pairs of samples, taking into account the larger uncertainty of the original GFAA measurements.

The agreement of laboratory measurements for the vast majority of the 349 wells sampled in 2000–2001 and 2003 is noteworthy for several reasons. The results demonstrate that the procedure followed to label the wells, relocate them, and sample and analyze the water was not flawed in any major way. The identification of specific wells would have been ambiguous if they had not been tagged, especially in those cases where relatives from the same cluster own more than one well. We have no explanation for the limited number of discrepancies that were observed and are in the process of retesting the wells and informing the owners of the results. No systematic temporal trend can be inferred from the data.

The number of outlier pairs indicating higher concentrations in 2003 ($n = 18$) is only slightly higher than the number of pairs with higher concentrations in 2001 ($n = 13$). When the 31 outliers are excluded, the differences in As concentrations determined by HR ICP-MS averages 2 ± 9 μg/L ($n = 318$) for samples collected in 2000–2001 and 2003. The limited amount of well-documented monitoring data available for Bangladesh aquifers suggest limited, if any, seasonal variations in groundwater As concentrations.
TABLE 1. Comparison of Laboratory and Field Kit Results for 799 Wells Tested by BAMWSP and a Partially Overlapping Set of 108 Wells that Were Retested by Our Team

<table>
<thead>
<tr>
<th>As concn (HR ICP-MS)</th>
<th>&lt;10 μg/L</th>
<th>10–50 μg/L</th>
<th>50–100 μg/L</th>
<th>&gt;100 μg/L</th>
<th>entire range</th>
</tr>
</thead>
<tbody>
<tr>
<td>no. of sampled wells with BAMWSP paint</td>
<td>236</td>
<td>187</td>
<td>139</td>
<td>237</td>
<td>799</td>
</tr>
<tr>
<td>no. of wells with incorrect result (%)</td>
<td>3 (1)</td>
<td>15 (8)</td>
<td>61 (44)</td>
<td>20 (8)</td>
<td>99 (12)</td>
</tr>
<tr>
<td>no. of wells retested in the field</td>
<td>28</td>
<td>18</td>
<td>43</td>
<td>19</td>
<td>108</td>
</tr>
<tr>
<td>no. of wells with incorrect result (%)</td>
<td>4 (14)</td>
<td>10 (56)</td>
<td>43 (100)</td>
<td>14 (74)</td>
<td>71 (66)</td>
</tr>
<tr>
<td>Hach 20 min</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>34 (79)</td>
<td>3 (16)</td>
<td>37 (34)</td>
</tr>
<tr>
<td>Hach 40 min</td>
<td>2 (7)</td>
<td>4 (22)</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td>7 (6)</td>
</tr>
</tbody>
</table>

* The laboratory data are taken as the reference to evaluate the field data.

FIGURE 3. Comparison of BAMWSP field tests with laboratory measurements for 799 randomly selected wells sampled in 2003. Each As concentration interval of 10 μg/L below 100 μg/L and of 50 μg/L from 100 to 300 μg/L is represented by at least 20 wells. The 300–1000 μg/L As concentration interval contains 44 wells. The shading emphasizes the area where the proportion of green wells does not agree with the laboratory data.

implication of the outcome of the resampling effort is that As measurements made in the field with a kit should in the vast majority of cases agree with the laboratory results, even if a particular well was tested in a different year or a different season.

Comparison of Laboratory and Field Data. In this analysis, we use HR ICP-MS data as the reference for evaluating the reliability of field tests with the Hach kit on the basis of which wells were painted red or green by BAMWSP workers (Table 1). Comparison of the results by the two methods shows that the spouts of 99 (12%) out of a total of 799 randomly selected wells were painted with the incorrect color. The largest proportion of errors (44%) is observed in the 50–100 μg/L range of As concentrations (n = 139). A different compilation of the results provides a better way of evaluating the actual threshold, if any, corresponding to the change in paint color applied by the BAWMSP workers (Figure 3). When the set of 799 BAMWSP results is subdivided into concentration intervals of 10 μg/L, the proportion of wells painted green gradually declines from ~95% at 30 μg/L As to ~15% at 100 μg/L As, rather than the sharp drop at 50 μg/L that was expected (Figure 3). These results should be robust since each of the concentration intervals is represented by at least 20 wells. Whereas the largest drop in the proportion of wells painted green is in the 50–60 μg/L range, the proportion of incorrectly classified wells is up to 15% even for As concentrations >100 μg/L (Figure 3).

The proportion of wells incorrectly classified by BAMWSP was particularly high (66%) for the subset of 108 wells selected by our team for retesting in the field (Table 1). To some extent, this reflects the selection of over half the wells on the basis of discrepancies relative to laboratory results. This is probably not the only explanation, however, since field tests by our team using a 20-min reaction time reduced the proportion of incorrect results for the same set of wells to 34%. More significantly, the proportion of incorrect results was reduced to 6% for the same set of wells by increasing the reaction time from 20 to 40 min (Table 1). The effect of increasing the reaction time was particularly dramatic in the 50–100 μg/L range of As concentrations, where the proportion of incorrect results was reduced from 79% to 2% (n = 43). The longer reaction time, however, also increased from 0 to 12% (n = 49) the proportion of wells containing <50 μg/L As that were incorrectly classified as unsafe relative to the Bangladesh standard. Analysis of an additional 30 inorganic constituents that included Na, Mg, Al, Si, P, S, K, Ca, Mn, Fe, Zn, and Sr (listing only those only present at concentrations >10 μg/L) did not reveal any systematic relationship between groundwater composition and whether laboratory and field data agreed or not (Supporting Information).

Discussion and Recommendations
Reliability of the Field Kit. The Hach kit, as deployed by BAMWSP workers in Araihazar in 2003, correctly classified 88% of wells relative to the Bangladesh standard for As in drinking water of 50 μg/L. Retesting of a subset of wells in the field by our team indicates that increasing the reaction time to 40 min is a modification of the procedure that could greatly increase measurement accuracy, which could be of particular benefit for the wells with concentrations between 50 and 100 μg/L As. Many of these wells had been incorrectly labeled as containing less than 50 μg/L As. The longer reaction time is desirable even if it is likely to discourage drinking or cooking with water containing 10–50 μg/L As since those wells do not meet the stricter WHO guideline for As in drinking water. Clearly, the Hach kit should continue to be used to test wells throughout Bangladesh and other countries affected by elevated As in groundwater.

Comparison of the BAMWSP results with our team’s Hach kit data also suggests that BAMWSP workers may not always have allowed the reaction to proceed for the prescribed 20 min in our study area (Table 1). This was confirmed by anecdotal reports of field workers feeling pressed for time because of the need to complete a certain number of tests within the day. Incentives may therefore be needed to reward good quality measurements in the field.

The outcome of this comparison of laboratory and field measurements is considerably more optimistic than that of similar studies conducted in Bangladesh (5) and West Bengal, India (6). One possible reason is that the three kits in ref 6 used a smaller sample volume (5–15 mL) than the Hach kit (50 mL). Our results clearly indicate that sending every tube well sample to a central laboratory for testing should not be a high priority, as was also pointed out in ref 11. The probably unintended consequence of such a recommendation issued...
in West Bengal has been that, because of the complex logistics involved in shipping samples and returning the results to the well owners, current government policy is to test only community wells while the vast majority of existing wells, which are private, remains untested.

Implications for Arsenic Mitigation. Considerable variations in the proportion of unsafe wells between neighboring villages have been pointed out by previous studies (3, 8). The prioritization of interventions should therefore be based on information disaggregated at the village level and not at the level of the upazila which typically covers several hundred villages (12). The comparison of BAMWSP data with our laboratory results indicates that field tests compiled for millions of wells provide a sound basis for ranking the severity of the situation faced by individual villages throughout the country.

Effective mitigation of the groundwater As crisis is urgently needed in Bangladesh and several other affected countries. Aside from its known association with various cancers and cardiovascular disease (13–15), new evidence shows that exposure to even low levels of As impairs the cognitive development of children (16). As a short-term measure, the sharing of existing wells could be promoted in villages where BAMWSP data indicate a significant proportion of existing safe wells (7). The installation of community wells appears to be the most viable option for many villages in the medium term (18). In Araihazar, we have sponsored the installation of nearly 50 community wells that tap deeper aquifers that are low in As. The community wells in Araihazar have become very popular and are each used, on average, by ~100 households living within a distance of 150 m (18).

Nearly all these wells draw their water from sandy deposits with a characteristic orange-brown color. Periodic monitoring shows that most of the community wells consistently meet the WHO guideline of 10 μg/L for As as well as the 500 μg/L guideline for Mn in drinking water (unpublished data, A.V.G.). The minimum drilling depth that was required to reach these aquifers varied from 50 to 200 m within the study area, however. This is a reflection of the variability of the local geology and supports the notion that the village is the appropriate scale to consider for mitigation (8, 18).

The sustained impact of both forms of interventions in our study area, the promotion of sharing of safe wells and the installation of community wells, has been confirmed by a dramatic drop of urinary As levels, a reliable indicator of exposure (unpublished data, J.I.G.). To our knowledge, the impact of a mitigation effort of this scale has not been documented for other approaches based on surface water treatment, As removal from groundwater, or rainwater collection. There are concerns about the long-term viability of deep wells, but these may have been exaggerated (3, 8, 19). In our study area, the groundwater withdrawal for personal use with hand-pumps is comparable to the recharge rate of the deeper aquifers tapped by the community wells (20). The much larger withdrawals by mechanized pumps for irrigation should be banned from using the deeper aquifers, however.

This study has shown that existing BAMWSP data based on field kits could in short order be used to prioritize and target the installation of community wells in thousands of affected villages. In parallel, it appears that a network of reliable testing services would have a major impact at the village level. In contrast to a national blanket testing campaign where many wells need to be tested in a day, the consequences of increasing the reaction time from 20 to 40 min to improve accuracy of routine testing at the local scale should be minimal. The installation of a community well to a safe depth is likely to result in the installation of private wells to the same depth. Because of the considerable geographic and depth variability of groundwater As concentrations, these wells will also need to be tested. If such a network can be created with help from the international community, then the parallel deployment of field kits for Mn should also be considered.

Acknowledgments

This work was supported by NIEHS Superfund Basic Research Program Grants NIH 1P42 ES105349 and NIH P30 ES 09089 and by the Earth Institute at Columbia University. We are grateful to the University Students for their effort in resampling the wells. The World Bank and the Hach Company were not involved in this study. This is Lamber-Doherty Earth Observatory Contribution No. 6676.

Supporting Information Available

Additional well data (Excel). This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited


(5) BRAC. Combating a Deadly Menace: Early Experiences with a Community-Based Arsenic Mitigation Project in Bangladesh; Research Monograph Series No. 16; Dhaka, Bangladesh, 2000.


(10) Gutzeit, H. Pharm. Zeitung 1879, 24, 263.


(15) Smith, A. H.; Goycoelea, M.; Haque, R.; Biggs, M. L. Marked increase in bladder and lung cancer mortality in a region of


Received for review June 14, 2004. Revised manuscript received October 6, 2004. Accepted October 7, 2004.

ES0491073