A statistical analysis of the pros and cons of field kits to guide well-switching in arsenic prone regions

Jameel Yusuf[1], Mozumder M. Rajib Hassan[2], van Geen Lex[3], and Harvey Charles F.[4]

[1]Massachusetts Institute of Technology
[2]Columbia University
[3]Lamont Doherty Earth Observatory
[4]MIT

November 16, 2022

Abstract

Field kits for testing the level of a toxicant in the environment are inherently less accurate than a laboratory instrument. Using a specific example, we argue here that kit measurements still have a key role to play when the spatial distribution of a toxicant is very heterogeneous. The context is provided by the groundwater arsenic problem in Bangladesh. We combine here two data sets, a blanket survey of 6595 wells over a 25 km^2 based on laboratory measurements and 900 paired kit and laboratory measurements from the same area. We explore different hypothetical mitigation scenarios based on actual data that rely on households with a high-arsenic well switching to a nearby low-arsenic well. We show that the decline in average exposure to arsenic from relying on kit rather than laboratory data is modest in relation to the logistical and financial challenge of delivering exclusively laboratory data. Our analysis indicates that the 50 ug/L threshold used in Bangladesh to distinguish safe and unsafe wells, rather than the WHO guideline of 10 ug/L, is close to optimal in terms of average exposure reduction. We also show, however, that providing kit data at the maximum possible resolution rather than merely classifying wells as unsafe or safe would be even better. These findings are relevant as the government of Bangladesh is about to launch a new blanket testing campaign of millions of wells using field kits.

A statistical analysis of the pros and cons of field kits to guide well-switching in arsenic prone regions

Yusuf Jameel[1], M. Rajib Mozumder[2,3], Alexander van Geen[2], and Charles F. Harvey[1]*

[1] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 15 Vassar Street, Cambridge, MA 02139.
[2] Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964.
[3]Now at Gradient, One Beacon St, Boston, MA 02108.

* corresponding author, Email: yjameel@mit.edu

Key Points:

1. Average reduction in drinking water arsenic exposure based on inaccurate field kit and accurate laboratory measurements are comparable.

2. Providing actual kit data rather than merely classifying wells as unsafe or safe would lead to a much higher reduction in arsenic exposure.

3. A threshold of 50 μg/L to distinguish safe and unsafe wells leads to higher reduction in arsenic exposure in comparison to a threshold of 10 μg/L.

Abstract

Field kits for testing the level of a toxicant in the environment are inherently less accurate than a laboratory instrument. Using a specific example, we argue here that kit measurements still have a key role to play when the spatial distribution of a toxicant is very heterogeneous. The context is provided by the groundwater arsenic problem in Bangladesh. We combine here two data sets, a blanket survey of 6595 wells over a 25 km^2 based on laboratory measurements and 900 paired kit and laboratory measurements from the same area. We explore different hypothetical mitigation scenarios based on actual data that rely on households with a high-arsenic well switching to a nearby low-arsenic well. We show that the decline in average exposure to arsenic from relying on kit rather than laboratory data is modest in relation to the logistical and financial challenge of delivering exclusively laboratory data. Our analysis indicates that the 50 μg/L threshold used in Bangladesh to distinguish safe and unsafe wells, rather than the WHO guideline of 10 μg/L, is close to optimal in terms of average exposure reduction. We also show, however, that providing kit data at the maximum possible resolution rather than merely classifying wells as unsafe or safe would be even better. These findings are relevant as the government of Bangladesh is about to launch a new blanket testing campaign of millions of wells using field kits.

1. Introduction

1.1. Background

In rural Bangladesh, and South Asia more generally, treated drinking water distributed through a piped system is rare. Sadly, millions drink arsenic-contaminated groundwater from their household well that adversely affect their health. An estimated 57 million people in Bangladesh are exposed to arsenic concentrations greater than 10 μg/L, the WHO standard, causing more than 100,000 excess spontaneous abortions and infant and adult deaths every year (Flanagan et al., 2012; Quansah et al., 2015).

Well testing conducted over a decade ago likely was the most successful arsenic mitigation program in Bangladesh by inducing millions of households to switch to a nearby safe well (Jamil et al., 2019; Van Geen et al., 2002). Implicitly, this success has been recognized by the government, along with the installation of deep low-arsenic wells, as it is about to launch a new wave of well testing across much of the country. The approach is viable because low-arsenic wells are often in close proximity to contaminated wells and households have shown willingness to make the extra effort of fetching water from another well and overcome social barriers to well sharing. Well-switching programs require measurements of arsenic concentration in most wells across a community but, unfortunately, most wells in Bangladesh are untested, partly due to the continuing installation of new wells (van Geen et al., 2014).

To fill this data gap, the government of Bangladesh has announced a several hundred-million-dollar project to test groundwater arsenic concentration across the country using inexpensive field kits. Field kits are used in Bangladesh instead of more accurate laboratory methods that require more resources: transport of the samples to labs, expensive spectrometers, and return of the results to well owners. Field kits can be performed on-site by local people with basic training. However, field kits are less accurate than spectrometric measurements conducted in laboratories and provide only categorical measurements representing nominal ranges (e.g. 50 to 100 μg/L). A recent comparative analysis between several types of kits concluded that improved precision and accuracy are necessary to employ kits for health-related decision making (Reddy et al., 2020). In contrast, it has been argued that existing kits have been effective in identifying arsenic contaminated wells (Ahmed et al., 2006; van Geen et al., 2005; George et al., 2012).

This paper focusses on two key questions concerning well-switching based on field kits. First, how and to what extent do the inaccuracies of a field kit diminish the effectiveness of well switching. Second, what arsenic concentration threshold should be used to identify the wells to avoid. In Bangladesh, a single threshold of 50 μg/L arsenic – which is higher than the WHO guideline of 10 μg/L - is used to categorize wells as either safe or unsafe. In 1993, when the WHO guideline for safe level of arsenic in drinking water was reduced to 10 μg/L , Bangladesh did not reduce its drinking water standard (Smith and Smith, 2004). To the best of our knowledge, no previous study has considered what concentration threshold might produce the best outcome in terms of well switching. This is important as the Bangladeshi government is spending millions of dollars on field test kits to implement well switching, and because millions of people may switch their water supply based on these measurements. Here, we compare the effectiveness of well switching recommendations based on accurate spectrometric-based and less accurate kit-based measurements using the actual distribution of arsenic in well-water documented in this particular region 20 years ago (van Geen et al., 2003). There is no reason to believe that this underlying distribution has changed drastically even under pumping conditions because groundwater arsenic levels are determined by exchange with a much larger pool of arsenic in the sediments (Fendorf et al., 2010). The spatial distribution of arsenic, lateral and vertical, in groundwater of this particularly region is also remarkable similar to that in the country overall (BGS/DPHE, 2001).

1.2. The Subtleties of Well-switching

The efficacy of well switching in terms of lowering exposure depends in surprising ways on spatial distribution of well arsenic, the threshold set for safe and unsafe wells, and the accuracy of testing methods. We illustrate some of these complexities with a hypothetical set of eight wells and their arsenic concentrations (Figure 1). The hypothetical example is meant to demonstrate the problems that motivate our analysis of a far larger data set. We consider two patterns of arsenic concentrations: one with concentrations that are spatially correlated so that there is a decreasing trend to the east; the other, with the same set of concentration values, but rearranged to be uncorrelated across the domain. Both of these examples use the same set of eight concentrations, just arranged differently, and hence create the same exposure before well switching, but produce very different exposures after people switch to nearby wells with lower arsenic concentrations (Panels A, B C and D in Figure 1). We evaluate the efficacy of the switching scenarios based on the net reduction of mean arsenic exposure.

First, we consider well switching for the ideal situation in which all arsenic concentrations are perfectly known and every participant switch to the well with the lowest arsenic concentration within 100 meters (panels A1 and A2). Considerable well-switching within that distance in response to testing has previously been documented (Chen et al., 2007; Madajewicz et al., 2007). This ideal, but unrealistic, switching scenario serves as an upper bound to which we compare more realistic scenarios with uncertain measurements and discrete thresholds. In the ideal case, where arsenic concentrations are spatially correlated and generally lower to the east (panel A1), well switching drops the mean concentration of consumed water from 134 to 38 μg/L. In this case, 3 of the 8 households “chain switch”, i.e. they switch to wells whose owners themselves switch to wells with even lower concentration, a behavior that may be unrealistic. When there is no spatial correlation (panel A2), there is no such chain switching and groups of households instead switch in clusters to a well with locally low concentrations, reducing the mean consumed concentration to 11 μg/L, an even lower value than the case where concentrations are spatially correlated.

Categorizing wells precludes switching within a category. Consequently, switching opportunities are lost when wells are categorized into “safe” above a threshold and “unsafe” below that threshold (Panels B and C). For example, when the threshold is 50 μg/L (panel B1), the wells with 652 and 167 μg/L do not switch to the well with 64 ug/L because all of three of these wells are in the “unsafe” category of above 50 μg/L. For this hypothetical arrangement of wells, a well switching program with a threshold value of 50 μg/L is therefore much less effective than the ideal case that uses continuous concentrations (Panel A1). The mean arsenic concentration is reduced only to 131 μg/L from 134, rather than 38 μg/L as in the case when continuous concentrations were used. There is also no chain switching when there are only two categories.

In the next example, we demonstrate that a different threshold value can lead to better or worse outcomes of well switching (Panels C1 and C2). In Panel C1, a threshold of 100 μg/L produces a much lower average consumed concentration than for a 50 μg/L threshold as in B1 because the well with 82 μg/L is now labeled safe and the two wells with the highest concentrations (652 and 167 μg/L) switch to this well. In contrast, a comparison of panels C2 and B2 shows how the outcome produced by 100 μg/L threshold can be worse than for a 50 μg/L threshold because the wells at 84 and 62 μg/L are now categorized as safe, and don't switch to wells with lower arsenic concentrations. The example proves that the optimal threshold value is not necessarily the concentration that has been deemed safe to drink based on health or other criteria. It raises an important question concerning well switching: what is the optimal threshold for categorizing wells as "safe" or "unsafe" to minimize arsenic exposure?

In the last part of this hypothetical example, we demonstrate how field-kit errors that lead to incorrect categorization also led to different well-switching outcomes. In panel D1, the mislabeling of a 64 μg/L well as safe, even though it is over the 50 μg/L threshold, surreptitiously leads to a better outcome. Because this well is mislabeled (due to an inaccurate field-kit measurement), neighboring households with even higher concentration wells now switch to it, and the overall mean consumed concentration falls from 134 μg/L to 50 μg/L, in fact much lower than the 131 μg/L for accurate measurements (Panel B1). In panel D2, the mislabeling of 47 μg/L well as unsafe leads to switching to a mislabeled well with higher concentration (64 μg/L) as safe. The mislabeling as well as the lack of spatial gradient leads to a higher reduction in kit-based arsenic switching from 134 μg/L to 23 μg/L in example D2. Perhaps counterintuitively, in this case, less accurate kit results lead to a greater reduction in arsenic exposure compared to accurate lab measurements.

1.3. Scope of this Analysis

In this paper, we analyze a large set of field data to statistically characterize outcomes for different well switching strategies and answer the questions demonstrated by the hypothetical examples above. We use arsenic concentrations measured in the laboratory across several thousand wells in Araihazar as our test data set. We supplement this with another set of field kit data from the same area paired with laboratory measurements. The Araihazar area has been the focus of numerous previous studies including behavioral studies that analyzed how household decisions to switch wells depend on information about well arsenic concentrations (Bennear et al., 2013; Huhmann et al., 2019; Madajewicz et al., 2007) and is the home to the extensive Health Effects of Arsenic Longitudinal Study (HEALS) longitudinal study health study (Ahsan et al., 2006). The site is also the locus of many geochemical studies of arsenic in groundwater, more specifically the vulnerability of low arsenic aquifers to contamination (Mihajlov et al., 2016; Mozumder et al., 2020). We focus here on questions that are important for designing a well-switching program: What are the probabilities of assigning correct (and incorrect) color placards to a well where arsenic concentrations are measured by field kits? In other words, how often do less accurate field-kit data lead to either a failure to correctly label a contaminated well (false negative) or mis-categorization of a safe well as unsafe (false positive)? How and to what extent does the inaccuracy of kit data diminish the effectiveness of well switching? What is the optimal threshold between "safe" and "unsafe" wells that minimizes exposure? How does the spatial pattern of arsenic concentrations impact the effectiveness of well switching?

2. Methods

2.1. Datasets

Two datasets previously collected as a part of the HEALS program in Araihazar Bangladesh provide the necessary data for our analysis. The first set pairs field kit measurements of arsenic concentration with accurate measurements made by inductively coupled plasma mass spectrometry (ICPMS) for 900 different wells (van Geen et al., 2014). This pairing enables a statistically assessment of errors in kit measurements. The ICPMS measurements have ±5% relative errors (Cheng et al., 2004). Field kit measurements are categorical, where each of nine different categories represents a range of arsenic concentration. We refer to these ranges as nominal because the actual concentrations can fall outside the range (Figure 2). The second dataset contains accurate measurements from 6595 wells representing every well within a 25 km^2 region that could be sampled in 2000-01 (van Geen et al., 2003). This large data set provides both the density of data to represent neighboring well concentrations and the extent of data to analyze a large-scale well switching program. The distribution of groundwater arsenic concertation in both the datasets are statistically similar.

2.2. Statistical Characterization of Field Kit Errors

We characterize the errors made in placing wells in safe and unsafe categories by analyzing the dataset of paired kit and spectrometric measurements from 900 wells in Araihazar (Figure 2). Government programs apply green or red paint to the pumphead to indicate if wells are safe or unsafe. However, in a kit-based testing of all of Araihazar upazila conducted in 2012, three colors were used for longer-lasting placards: blue for arsenic <10 μg/L, green for arsenic between 10 and 50 μg/L, and red for wells with arsenic >50 μg/L.

To calculate the probabilities of an incorrect label, we first estimate a probability density function $f_n(\theta)$ for the actual concentrations (θ)within each kit category n. We take the laboratory-measured concentrations associated with each kit category and fit a parametric probability distribution function (Figure 3). We then use these nine (for each of the kit category) probability density functions to calculate the conditional probabilities of assigning a particular category conditioned on the spectrometric arsenic measurements:

$$P(Kit\ category\ = n|\ Arsenic\ =\ \theta)\ = \frac{f_n(\theta)}{\sum_1^9 f_i(\theta)} \quad (1)$$

Where the LHS in equation 1 provides the probability of observing each kit category if the laboratory measured concentration is θ and i corresponds to the nine nominal kit categories. Table 1 provides a worked-out example for calculating the conditional probabilities of the different kit categories for a well with arsenic concentration of 100 μg/L.

These conditional probability density functions are then used to quantify the probability that a kit measurement falls in any category, correct or incorrect, given a spectroscopic (laboratory) measurement, and hence errors that occur when kit measurements are used to label wells. The probability of assigning blue placards to wells with arsenic between 0 and 10 ug/L (correct assignment) is:

$$P(i = \{1,2\}|\theta < 10)\ = \frac{\sum_1^2 \int_0^{10}\ f_i(\theta)d\theta}{\sum_1^9 \int_0^{10}\ f_i(\theta)d\theta} \quad (2)$$

Where i =1 and 2 corresponds with nominal kit range of 0 -1 ug/L and 1-10 ug/L. Similarly, the probability of assigning green or red placards when the accurate measurement is less than 10 ug/L (incorrect assignment) is given by:

$$P(i = \{3,4,5,6,7,8,9\}|\theta < 10) = \frac{\sum_3^9 \int_0^{10} f_i(\theta)d\theta}{\sum_1^9 \int_0^{10} f_i(\theta)d\theta} \quad (3)$$

Where i = 3 to 9 corresponds to nominal kit categories with range of >10 ug/L (Figure 2).

For wells between 10 and 50 ug/L, the probability of assigning correct (green), false negative (blue) and false positive (red) placard is given by equations 4 to 6 respectively:

$$P(i = \{3,4\}|10 < \theta < 50) = \frac{\sum_3^4 \int_{10}^{50} f_i(\theta)d\theta}{\sum_1^9 \int_{10}^{50} f_i(\theta)d\theta} \quad (4)$$

$$P(i = \{1,2\}|10 < \theta < 50) = \frac{\sum_1^2 \int_{10}^{50} f_i(\theta)d\theta}{\sum_1^9 \int_{10}^{50} f_i(\theta)d\theta} \quad (5)$$

$$P(i = \{5,6,7,8,9\}|10 < \theta < 50) = \frac{\sum_5^9 \int_{10}^{50} f_i(\theta)d\theta}{\sum_1^9 \int_{10}^{50} f_i(\theta)d\theta} \quad (6)$$

For wells >50 ug/L, the conditional probability of correct (red) and false negative (green or blue) assignments is given by equations 7 and 8 respectively:

$$P(i = \{5,6,7,8,9\}|\theta > 50) = \frac{\sum_5^9 \int_{50}^{\infty} f_i(\theta)d\theta}{\sum_1^9 \int_{50}^{\infty} f_i(\theta)d\theta} \quad (7)$$

$$P(i = \{1,2,3,4\}|\theta > 50) = \frac{\sum_1^4 \int_{50}^{\infty} f_i(\theta)d\theta}{\sum_1^9 \int_{50}^{\infty} f_i(\theta)d\theta} \quad (8)$$

We also calculated the conditional probability of a well being assigned as safe and unsafe as following:

$$P(i = \{1,2,3,4\}|\ Arsenic = \theta) = \frac{\sum_1^4 f_i(\theta)d\theta}{\sum_1^9 f_i(\theta)d\theta} \quad (9)$$

$$P(i = \{5,6,7,8,9\}|Arsenic = \theta) = 1 - P(i = \{1,2,3,4\}|\ Arsenic = \theta) \quad (10)$$

2.3. Well Switching from blanket testing:

To investigate the efficacy of well-switching based on different criteria we use the large data set of accurate arsenic measurements that represents nearly all wells in a 25 km^2 portion of Araihazar. This data set does not contain field-kit measurements but, because we have analyzed the paired data set and calculated the pdf (probability distribution function) of true arsenic concentration for each kit, we can simulate field-kit measurements from the accurate measurements (see Table 1).

We simulate switching for all wells within 100 m of each other – i.e. a household will switch to a better well if it is within 100 m. In Bangladesh and India, the probability of well switching drops with distance to a well (Barnwal et al., 2017; Gelman et al., 2004; Madajewicz et al., 2007; Pattanayak and Pfaff, 2009), so that the probability of switching is low (<0.3) if the distance between the unsafe and the safe well is greater than 100 m. In our analysis, everyone switches from an unsafe well if a safe well is within 100 meters.

We evaluated eleven different switching plans based on laboratory-measured arsenic concentrations and the simulated kit categories obtained using the statistical relationship established between the nominal field kit categories and the actual arsenic concentrations (section 2.2.). We judged the effectiveness of each plan by calculating the mean reduction in arsenic exposure pre-and post-switching. We divide the eleven plans into 3 groups. In Group A, we investigate switching based upon spectrometric measurements. In Group B, switching is based on the simulated kit categories. In Group C, we investigate the effects of spatial correlation on switching.

If groundwater arsenic is measured accurately, the only switches will be to less contaminated wells. However, categorizing of wells as safe or unsafe using inaccurate field kits leads to a variety of poor switches. First, a switch can take place between a correctly identified contaminated well (nominal kit categories 5 to 9) to an incorrectly identified safe well (bad switching). Second, a switch can take place between a safe well that is incorrectly identified as contaminated to another safe well that is correctly identified as safe (unnecessary switching). Third, a switch can take place between a safe well incorrectly identified as unsafe to a contaminated well that is incorrectly identified as safe (very bad switching). Fourth, an unsafe well incorrectly assigned as safe will not switch (missed opportunity for switching). All these possible switching scenarios and their associated probabilities are provided in Table 2. Thus, switching based on field kits can result in unnecessary, bad or failed switching which is not the case for switching based on laboratory measurements.

2.3.1. Group A: Switching based on laboratory measurements

Scenario A1: Everyone switches to the well with the lowest arsenic concentration within a 100 m radius. This ideal but unrealistic plan serves as a point of comparison for more realistic scenarios and provides an upper bound on the possible reduction in exposure.

Scenario A2: We investigate the effects of the modest errors in laboratory measurements of ± 10%. This scenario is similar to scenario 1, except that we add a random normal error with a standard deviation ±10% the value to each data point. The purpose of this scenario is to quantify the effect of analytical uncertainty on the well switching exercise. We did not add uncertainty to wells with arsenic concentration of 0 μg/L and the concentration of these wells kept at 0 μg/L – primarily because several wells with a measured concentration will have an unrealistic negative concentration after incorporating the uncertainty.

Scenario A3: We consider the effect of labeling wells as categorically safe or unsafe rather than using concentrations. Thus, wells with arsenic concentration >50 μg/L and <50 μg/L were labeled red and green respectively. Everyone using the red well switches to a green well – if such a well exists within a 100 m radius.

Scenario A4: We use the three categories, as is the recent practice in Araihazar where the wells were categorized in three categories (van Geen et al., 2014) instead of the two categories used elsewhere in Bangladesh. In Araihazar, wells with arsenic <10 μg/L are labeled blue and wells with arsenic between 10 and 50 μg/L are labeled green, and above 50 are labeled red. Consumers using red wells switch to the nearest blue well (if any was present) in the 100 m radius of the well. If there were no blue well consumers switched to the nearest green wells. If there was neither a blue nor a green well in the 100 m radius, the consumers did not switch.

Scenario A5: Here we find the optimal switching concentration such that the mean exposure after switching is the lowest. The decision to label wells >50 ug/L as contaminated and wells with concertation <50 μg/L as uncontaminated in Bangladesh was not chosen specifically to optimize health outcomes. For instance, wells with arsenic concertation just below 50 μg/L (such 45 μg/L) cannot switch to a nearby well with lower arsenic as both would be labeled green. Similarly, in case where there is no safe well in the vicinity of a contaminated well (such as a well with 230 μg/L), the well cannot switch to a nearby less contaminated well (such as a well with arsenic concentration of 60 μg/L) as both would be labeled red and based upon the color it would be impossible for the consumers to know which well is more contaminated and vie-versa. We investigated the arsenic concertation (10 -100 ppb, with an increment of 1 ppb) below and above which a well is labeled safe (green) and unsafe (red) to find the switching concentration above which the wells should be labeled red and below which it should be labeled green such that the mean exposure post-switching (based on well labels) is minimum.

2.3.2. Group B: Switching based on kit measurements

In this group of well-switching simulations, we consider well switching plans based on kit measurements of arsenic concentration by simulating kit measurements.

Scenario B1: Here we consider well switching based on all nine kit categories and using the statistics of categorization errors (Section 2.1.) to simulate mis-categorizations. Consumers of each well switched to the well assigned with the lowest kit category within a 100 m radius. This plan differs from typical plans that use only two categories.

Scenario B2: Here we consider the typical approach of labeling wells in only two categories, safe and unsafe. Wells with kit categories of 5 and above (i.e. nominal arsenic range of 50 - 100 μg/L and above) were labeled red and wells with kit categories of 1 to 4 (i.e. nominal arsenic range of less than 50 μg/L) were labeled green. Consumers of the red wells switch to the nearest green well (if any such well was present within a 100 m radius). This is analogous to switching scenario A3 based on laboratory measurements and represents the commonly practiced switching scenario in Bangladesh.

Scenario B3: Wells were labeled in three colors as has been done in Araihazar (analogous to spectrometric based switching scenario A4). Wells with kit categories of 0 and 1 were labeled blue, categories 3 and 4 were labeled green and categories 5 and above labeled red. We then assigned residents using red wells to switch to the nearest blue well (if any present) or else switch to the nearest green wells within a 100 m radius. If there is no blue or green labeled well, residents do not switch.

Scenario B4: Here we find the category above which the wells should be labeled red and below which it should be labeled green such that the mean exposure post switching is the lowest. Recent practice in Bangladesh has been to label wells that falls in categories 5 to 9 red. Here we consider whether this is the optimal threshold for reducing mean exposure post-switching. For this exercise we evaluated the exposure post well switching for all the 9 categories below and above which the well are labeled as safe and unsafe. For example, we compare the scenario when categories 2 to 9 are labeled as unsafe with the scenario when categories 6 to 9 are labeled as unsafe.

<u>2.3.3. Group C: Effects of spatial correlation in arsenic concentrations.</u>

In this set of simulations, we consider how the efficacy of well switching is affected by spatial correlations in arsenic concentrations across wells. Where arsenic concentrations are spatially correlated, well switching is limited because contaminated wells are more likely surrounded by contaminated wells and safe wells are surrounded by safe wells. Hence, the possibility of switching depends not only on the identification of safe and unsafe wells but also on spatial pattern of well arsenic concentrations.

Distribution of arsenic in Araihazar are weakly spatially corelated at small scales and contain some larger scale features, particularly a large cluster of low arsenic wells in the northwestern part of the district (Figure 4). To investigate the impact of these patterns on the effectiveness of well switching, we applied 2 hypothetical switching scenarios that removed spatial correlation. We randomly reassigned each well an arsenic concentration (and the corresponding simulated kit category) from the distribution of the 6595 wells (without replacement). Subsequently, we simulated well switching based upon the reassigned arsenic concentration to each well.

<u>Scenario C1</u>: Here we consider the effect of spatial correlation when measurements are accurate. This is analogous to scenario A1 except that the distribution of arsenic is not spatially corelated.

<u>Scenario C2</u>: Finally, we consider here the effect of spatial correlation when kit measurements are used. Analogous to scenario B1, we simulate switching based upon the reassigned simulated kit categories except that the distribution of arsenic is not spatially corelated.

3. Results

<u>3.1. Impact of uncertainty on assignment of wells to safe and unsafe categories</u>

Based on Kolmogorov-Smirnov (K-S) tests, a gamma distribution was the best fitting parametric function for the accurate arsenic measurements within each kit categories, except for categories 1 and 2. These data were fitted with Weibull and exponential distributions, respectively (Figure 3). For kit categories 6 to 9, the normal distribution was also a good fit to the accurate measurements, however we chose the gamma distribution as it is positively defined.

The estimate probabilities of assigning wells to different categories (with the corresponding color placards) are summarized in Figure 5. Mis-categorization of wells is most likely where the arsenic concentration is close to the threshold, and the probability of error falls off rapidly for concentrations that are far from thresholds. Thus, extremely contaminated wells are unlikely to be classified as safe by the field kits. For example, for arsenic concentrations above 200 μg/L, the probability of incorrectly assigning a well to safe, blue or green placard, was very small (<0.001).

<u>3.2. Well switching based on accurate arsenic data (group A)</u>

The first group of scenarios (group A) contains simulations based on accurate measurements (Table 3). Scenario A1 is the ideal base case: switching is based on continuous accurate arsenic data. In this case, 84% of consumers reduce their arsenic exposure by switching and the mean arsenic exposure of the residents decreased from 134 μg/L pre-switching to 17 μg/L after switching. Scenario A2 investigates the impact of the analytical uncertainty in accurate laboratory measurements of arsenic on well-switching. The reduction in arsenic exposure was similar to A1 suggesting that analytical uncertainty in laboratory measurements has a negligible influence on the outcome of well switching.

Scenarios A3 and A4 investigate the impact of categorizing wells based an accurate measurement. After sorting wells into two categories, red wells (As >50 μg/L) and green wells (As < 50 μg/L), only 43% of the residents lowered their arsenic concentration and the mean exposure post well switching was 37 μg/L. Fewer switches occurred than in scenario A1 because no switching occurs between wells with As < 50 μg/L. In scenario A4 wells were labeled in three categories and again 43% of the residents switched to lower arsenic concentration wells and the mean exposure post well switching was 35 μg/L – slightly lower than scenario A3. A comparison of scenarios A3 and A4 suggests that the fraction of residents that lower their arsenic exposure is similar when the wells are grouped in 2 or 3 color categories, however, the net reduction in arsenic exposure for 3 groups is slightly better because consumers can switch to wells with low arsenic (<10 μg/L, blue wells) where possible. Since wells with As <10 μg/L are mostly concentrated in the northwestern part of Araihazar, the decrease in arsenic exposure by labeling the wells in three categories was minimal; however, if the wells with As < 10 μg/L were truly randomly distributed the decrease would have been higher (see scenarios C1 and C2).

The optimal threshold concentration that minimizes mean exposure (scenario A5) is 41μg/L (Figure 6a), producing a post-switching mean exposure of 35 μg/L. This value is only 1 μg/L lower than the mean exposure at the 50 ug/L cutoff that is currently used in Bangladesh (scenario A3).

3.3. Well switching based on simulated kit categories (group B)

In this group of scenarios, we use simulated kit measurements to assess the impact of kit measurement errors. Unlike accurate measurements, kit measurements lead to some switches from lower to higher arsenic concentrations (Table 3, second row from the bottom). Scenario B1 describes the results when all kit categories are used to label wells. With these nine categorical labels, 69% of the consumers reduce their arsenic exposure and 25% of consumers keep the same level of exposure. Because of the inaccuracies of kit measurements, exposure increases for 6% of the consumers. The mean arsenic exposure of the residents was 25 μg/L after well switching.

Scenario B2 describes the typical case across Bangladesh: kit measurements are used to categorize wells as safe or unsafe with a nominal threshold of 50 μg/L. In this case fewer (42%) of the residents lower their arsenic concentration, 54% of residents have the same level of exposure and 3.5% of the residents increase their exposure. The mean exposure post well switching was 41 μg/L, still a big reduction from the average across wells of 134 μg/L.

Scenario B3 considers the atypical approach used in Araihazar: Three categories of wells (blue, green and red), rather than just two. The fraction of residents who reduced their exposure was higher than scenario B2 (Table 3) and the exposure was reduced to 35 μg/L, about 6 μg/L lower than with two categories, scenario B3.

The optimal threshold category for switching is the same as what is currently used in Bangladesh to assign the wells green and red placards, kit categories 1 through 4 are labeled green and wells with kit categories 5 and above are labeled red (Figure 6b).

3.4. Well switching based on randomization of arsenic concentration (group C)

The Group C scenarios both consider the effects of spatial correlation across wells on well switching by erasing this correlation. Scenario C1 considers the case of accurate measurements and C2 considers kit measurements. In both scenarios well switching becomes extremely effective: the exposure drops to 6 μg/L and 8 μg/L when arsenic concentrations are random in space because many more households have neighboring wells to switch to (84% and 76%, Table 3).

4. Discussion

4.1. Is well switching useful?

Exposure to high level of arsenic in drinking water is still pervasive in South Asia with more than 40 million people exposed to high level of arsenic in drinking water in Bangladesh alone (Jamil et al., 2019). Several strategies have been proposed in the past two decades to reduce arsenic exposure in drinking water including filtration of pond and surface water, removal of arsenic at the household level using purification filters, community filtration systems, rainwater harvesting and well switching (Ahmed et al., 2006). However, all the above methods except well switching have been deemed unsustainable or expensive due to multiple reasons. Sand filtration is unsuitable due to high concentration of fecal contaminants in surface water (Howard et al. 2006) and the inability of these filters to remove them. Similarly, high cost of household filters and regular maintenance of community filtration processes has led to their limited success in reducing population level arsenic exposure (Krupoff et al., 2020). A recent analysis by Jamil et al., (2019) suggests that well testing and subsequent switching leads to the largest decrease in population level arsenic exposure and is economically the most viable solution for reducing population level arsenic exposure. The total cost for well testing ($1) and subsequent switching is significantly lower than the cost per person associated with installing low arsenic deep wells ($143) and operating a treated piped water supply system ($158).

Although well testing based on kit measurement appears to be economically feasible and logistically viable, a major shortcoming of well testing that has been claimed is its lack of accuracy and precision (Jakariya et al., 2007; Reddy et al., 2020). Laboratory testing per well can range between $6-$30 and can increase the cost of well testing multifold, thereby greatly reducing the economic benefits provided by field kits in regard to other mitigation techniques. An important question to ask is what are the pros and cons of using field kits for well switching, how do they compare with laboratory-based measurements for well switching and at a community level it is a sustainable option to recommend for large scale well switching?

For all the switching scenarios presented here, the mean exposure post switching was substantially lower than the arsenic exposure pre-switching (Table 3). Excluding the hypothetical scenarios where the spatial distribution of well-water arsenic was randomized (scenarios C1 and C2), the net arsenic exposure post switching in Araihazar was 3 to 7 times lower than the exposure before switching. Even for the simulated kit-based switching scenarios, the net decrease in arsenic exposure was 3 to 5 times lower. The most important outcome of well switching (in all scenarios) was the ability to reduce the exposure of consumers using highly contaminated wells (> 100 μg/L) to significantly lower levels. Even for least effective scenario (B2), the average exposure for people exposed to wells with arsenic > 100 μg/L reduced from 214 μg/L to 74 μg/L.

The analysis presented here complements the economic analysis by Jamil et al., (2019) and provides support that well switching based on kit measurements is not only economically feasible but it also very effective in reducing population level arsenic exposure. Therefore, even though kit measurements can be inaccurate, they can lead to significant decrease in arsenic exposure at community level. The higher accuracy laboratory measurements render them only marginally better in terms of reducing arsenic exposure.

Notwithstanding the net decrease in arsenic following laboratory and kit-based measurements, it has been well documented that the well switching is not complete due to multiple factors including the distance between safe and unsafe wells and socioeconomic factors. In India, Barnwal et al. (2017) showed that the probability of well switching decreases rapidly as the distance between the safe and the unsafe well increases – if the distance between the safe and unsafe well is <10 m the probability of well switching is ~ 0.4 and if the distance is >100 m, the probability is <0.25. In Bangladesh, Madajewicz et al., (2007) reported that in Araihazar, 60% of the people who realized they were using contaminated well switched to a safe well within 1 year. These are high response levels, even if the maximum level of exposure reduction was not achieved and household knew that they were using contaminated water. One reason may be that many households with a low arsenic well might not be willing to share their wells with their neighbors; households with lower socioeconomic status find it more challenging to switch to safe wells (Madajewicz et al., 2007). This is why we assumed that only 50% of the household switches after realizing that they are using contaminated wells. The implication is that there is considerable potential for additional switching and more attention should be paid to ways of encouraging well switching and sharing among neighbors.

4.2. Comparing laboratory- and kit-based switching

The arsenic exposure post well switching and the proportion of households managing to switch were similar based on laboratory and kit measurements (Table 3). There are three major factors behind this surprisingly good outcome for kit-based switching. First is the ability of the kits to correctly identify the uncontaminated (arsenic $\leq$10 μg/L) and highly contaminated wells (>100 μg/L) with a high degree of accuracy (> 0.95, Figure 5a). Second, the distribution of groundwater arsenic in Araihazar is non-normal. On the order of 30% of wells contain <10 μg/L arsenic, 34% of the wells contain between 10 and 100 μg/L, and 36% contain > 100 μg/L arsenic. This resulted in the accurate labeling by the kit for approximately 66% of the wells. Indeed, the majority of mis-categorizations (false positive and false negative) was observed for the remaining 34% of the wells with true concentration between 10 and 100 μg/L. Third the overall degree of spatial autocorrelation in the distribution of groundwater arsenic was low – although for some pockets arsenic concentration were strongly corelated (Figure 4). Therefore, contaminated wells and uncontaminated or lower contaminated wells (Figure 4) were always in close proximity resulting in large number of switching.

The comparative analysis also highlights three major limitations of switching based on kit measurements. First, is the continued exposure of consumers using contaminated wells that were incorrectly assigned as safe. More than 22% of the wells with arsenic between 50 and 75 ug/L were assigned a kit category between 1 and 4 (i.e. blue or green placard). This prevented them from switching to a nearby safe (or less contaminated) well. Secondly, more than 40% of the wells with arsenic between 20 ug/L and 50 ug/L were incorrectly assigned kit category of 5 and above (i.e. red placard). This resulted in unnecessary switching by consumers using these wells.

In general, most consumers switched to a correctly categorized nearby uncontaminated well, however, some consumers switched to a well with higher arsenic concentration than their original well (bad switching). This led to an increase in arsenic exposure for 3.5% and 2% of the consumers in scenarios B2 and B3 respectively. Although this switching is extremely undesirable, the net increase in arsenic exposure of consumers experiencing bad switching was not high (15 ug/L for scenario B2).

4.3. Should wells be grouped in 3 color categories?

Typically, well switching exercise are based on the color of the placards placed on the well. Across Bangladesh, wells have been labeled green (<50 μg/L) or red (>50 μg/L), however recently in the Araihazar district wells were labeled blue (<10 μg/L), green (10-50 μg/L) or red (>50 μg/L). It is important to ask if there is an added advantage in labeling wells in three color categories. Our analysis suggests that the proportion of population switching from contaminated well to uncontaminated well based on two- or three-color placards are the same (compare scenarios A3 and A4 and B2 and B3, Table 3). Based on laboratory measurements the mean exposure post switching is comparable for both the scenarios (34 and 36.6 μg/L based on two- and three-color placard categories respectively). The difference based on the simulated kit categories was slightly higher (mean exposure post switching was 35 and 41 μg/L based on two- and three-color placard categories respectively) but not very large. In Araihazar, it appears that grouping the wells in 3 categories is only slightly more beneficial – the main factor driving this pattern is the clustering of majority of the wells with arsenic < 10 μg/L (i.e. blue wells) in the northwestern part of the district. Therefore, the benefit provided by grouping the wells in 3 color categories only helped a minority of the consumers.

It is worth noting that the mean reduction in exposure based on actual concertation (and actual kit categories) was almost twice as low than those based on color categories (compare scenarios A1 and A3 and B1 and B2 respectively, Table 3). Thus, if placards placed on the well also included the concentration (or the kit categories when field kits are used), the possible reduction in exposure could be substantially higher. In the original HEALS study, actual arsenic concentration was included on the well placard (Chen et al., 2007) and 58% of the 6,512 participants using contaminated wells (As ≥50 μg/L) switched to other wells. Therefore, including arsenic concertation (or the kit category) on the color placard that might cost an additional $1.5 but could lead to higher switching rates. Since well-switching is voluntary, providing the actual concertation (or the kit categories) would also provide the consumers more freedom in deciding if they want to switch and which well to switch to.

According to the Bangladesh Arsenic Mitigation and Water Supply Program (BAMWSP) survey of 2000-05, the number of wells with As > 50 μg/L in Araihazar and across Bangladesh are comparable (29% and 32% respectively). Additionally, the spatial heterogeneity – which is critical for effective well switching – across the upazilas (sub-districts) in Bangladesh is also comparable to Araihazar (Jamil et al., 2019). Although the testing under BAMWSP underestimated the number of high As wells (van Geen et al., 2005), the similarity in the number of well with As > 50 μg/L and in the spatial heterogeneity of groundwater arsenic concentration suggest that the findings presented here are relevant across the country. The number of wells have increased rapidly in Bangladesh in the last 10 years (Jamil et al., 2019), and most of these wells are untested for As. After the end of BAMWSP camping in 2005 there has been no blanket testing in Bangladesh and the current nationwide arsenic exposure in Bangladesh is unquantified.

Our analysis suggests that a nationwide blanket testing followed by widescale well switching has the potential to reduce mean arsenic exposure to concentrations lower than the current Bangladesh standard of 50 μg/L for most of the districts.

4.4. Is 50 ug/L the optimal level for labeling unsafe wells?

From a health perspective, various drinking water standards or the WHO guideline for arsenic are somewhat arbitrary; the WHO guideline of 10 μg/L is most widely referred to globally (Ahmad and Bhattacharya, 2019), however, standard for arsenic in drinking water also vary regionally. In the Netherlands, for instance, the voluntarily target of arsenic in drinking water is <1 μg/L (Ahmad et al., 2020). In the US, the EPA lowered the drinking water standard for arsenic from 50 to 10 μg/L as recently as 2001, but the state of New Jersey has lowered it to 5 ug/L. In Bangladesh and Pakistan, and until recently in India, the standard for arsenic in drinking water is still 50 μg/L. There is a continuum in toxicity across the range of arsenic concentrations and health effects do not suddenly appear with an increase from 9 to 11 μg/L or from 45 to 55 μg/L. Therefore, wells with arsenic >50 μg/L and/or wells with observed kit categories of 5 to 9 are labeled red (i.e. contaminated) and the well with arsenic < 50 μg/L are labeled green (i.e. safe). Consequently, users of wells with > 50 μg/L arsenic (i.e, observed field kit categories of 5 to 9) have been encouraged by the Bangladeshi government to switch to the nearby safe wells.

To our knowledge, the safe threshold of 50 μg/L in Bangladesh was not chosen to minimize arsenic exposure post well switching. As discussed in section 1.2, the optimal threshold values for well switching are not necessarily the concentration that has been deemed safe to drink and the optimal switching concentration can vary from region to region. Our analysis shows that the maximum reduction is arsenic exposure is observed if wells with As > 41 μg/L and wells with observed kit categories of 5 to 9 are considered as unsafe (Figure 6). This suggest that the currently used criteria of switching consumers using wells with arsenic > 50 μg/L (and observed kit categories of 5 to 9) happens to be fairly close to the optimal switching concentration. For switching based on 50 μg/L threshold, the mean exposure was always lower than 41 μg/L (scenario B2). In comparison, if all the wells with arsenic >10 μg/L are labeled as contaminated the mean exposure post switching would be 46 μg/L or greater. Similarly, labeling wells with simulated kit category of 2 to 9 as contaminated would result in a mean exposure 48 μg/L after well switching. Therefore, from a switching perspective, labeling well with arsenic > 50 μg/L as unsafe would lead to much higher reduction in exposure post switching than labeling wells with arsenic >10 μg/L as unsafe.

4.5. Are kit tests preferable to laboratory measurements?

In one way, the higher accuracy of laboratory measurements is preferable for identifying the status of a well with respect to arsenic. However, kit-based results that could be imprecise (Reddy et al., 2020) should be interpreted in a broader framework where a major motivating factor in using the less accurate kit-based measurements are their cost-effectiveness, rapid throughput and independence from expensive spectrometric instruments that are often unavailable in developing and lower income countries. Funds available at sub-district and village level are often limited; therefore, we evaluated if a limited amount of fund is available, whether it is more fruitful to test small number of well using the more accurate but expensive spectrometric measurements or test large number of wells using kits.

We address this question as a case study, assuming that a village is allocated $2000 to test arsenic in the groundwater wells. With that amount, the village can accurately measure arsenic in 200 wells using in the laboratory ($10 per sample, (Gelman et al., 2004)) or measure 2000 wells albeit with less accuracy using field kit ($1 per sample (Ahmed et al., 2006)). If we assume that the distribution of arsenic concentration in this village is similar to Araihazar, then kit measurements could lead to a possible reduction of arsenic exposure for 6000 consumers (assuming 10 consumers per well) from >50 μg/L to <50 μg/L. However, 140 people would most likely experience an increase in arsenic exposure due to the uncertainties associated with kit measurements. In contrast, laboratory measurements would lead to a decrease in arsenic exposure for 800 people from >50 μg/L to <50 μg/L with no one experiencing an increase in exposure. Thus, with a limited budget, kit measurements can reduce arsenic exposure for almost 8 times more people than laboratory measurements. However, this also led to an increased exposure of approximately 15 out of every 1000 people to higher levels of arsenic. This does lead to a moral dilemma from having to choose between a laboratory method that helps a small fraction of the population but does not adversely affect a single person in the population and a field-based method that helps a larger proportion of the population but could increase the arsenic exposure of a small proportion of the population.

From a utilitarian perspective that considers benefits to the population overall, the second scenario is clearly preferable. If increased arsenic exposures for a small proportion of the population is acceptable, then the case study provides compelling evidence that large quantities of lower-grade and imperfect kit-based measurements may be more effective in mitigating arsenic exposure than a small number of more precise spectrometric measurements – at least for places with arsenic distribution similar to Araihazar.

4.6. Promoting well-testing to mitigate arsenic exposure

It is estimated that arsenic related mortality would cost Bangladesh almost $12.5 billion in the next 20 years (Flanagan et al., 2012) and reducing arsenic in drinking water arsenic is an important goal of the Bangladesh government. Currently, well switching is the most cost-effective and is a scalable solution for reducing arsenic in drinking water across Bangladesh. Our analysis suggests that imprecise kit measurements can reduce arsenic exposure of more than 85% of the population of Araihazar to <50 μg/L (Bangladesh standard) if well switching was complete. However, well switching rate is moderate (varies between 30% to 60%) due to multiple factors (Barnwal et al., 2017; Jamil et al., 2019; Madajewicz et al., 2007) and the effective exposure at the population level remains high.

Large-scale decrease in arsenic exposure can only be achieved if wells are extensively tested and the results are shared with the households. Selling tests is not an option as it has already been shown that the demand for a field-kit test drop rapidly at any price that could potentially sustain a commercial testing service (Barnwal et al., 2017; Tarozzi et al., 2020). Therefore, testing should be offered free and the results should be shared with all households.

Krupoff et al., (2020) provided an analysis of well switching in Bangladesh from the perspective of the social sciences and suggested multiple reasons for modest switching rates and provided recommendations for increasing well switching rates in Bangladesh. First, the low rates of well switching could be the failure to provide the information to the consumers. Tests are commonly conducted by representatives who leave the village after performing the test leaving little opportunity to reinforce the information. In this regard training community members to perform arsenic measurements locally and constantly reinforce the information might be more help promote well switching. Providing monetary compensations might increase the commitment from the community members involved in testing and promoting well switching (BenYishay and Mobarak, 2019). Also important would be to develop a mechanism that promotes well sharing – such as combining testing with a community commitment (Inauen et al., 2014).

<u>4.7. Low spatial autocorrelation is essential for effective well switching</u>

The degree of spatial autocorrelation in arsenic concentration of the 6595 wells in Araihazar is low (Moran's I = 0.1, $p< 0.05$), however there is a large cluster of well with arsenic <10 ug/L in the northwest region and arsenic >50 μg/L in the southwest region. After randomizing the arsenic concentration in the well (Moran's I = -0.0007, p=0.35 after randomization), the mean exposure post switching decreased to 6 ug/L (using accurate spectrometric data) and 7.5 μg/L (using simulated kit categories, Table 3). The large decrease after randomization suggests that vast majority of the wells managed to switch to a blue well (arsenic <10 μg/L) in their vicinity. This highlights the importance of spatial autocorrelation in well switching exercise – in Araihazar even though the degree of spatial correlation is low, yet several consumers (around 15%) were unable to switch due to lack of uncontaminated wells in their 100 m radius. In villages where groundwater arsenic concertation is strongly spatially autocorrelated, the effectiveness of well switching would be fairly limited, however if the spatial autocorrelation in arsenic concertation is low well switching exercise would be fairly effective. Across much of Bangladesh, spatial correlation in groundwater arsenic is low (Gelman et al., 2004; Yu et al., 2003) providing strength to well switching as an effective approach to reducing arsenic exposure in drinking water.

5. Conclusions

The number of groundwater wells in Bangladesh has increased steadily (Dey et al., 2017; Jamil et al., 2019) and on the order of >1 million wells/year continue to be installed (van Geen et al., 2014). In most villages, only a small minority of wells are tested (George et al., 2012; Jamil et al., 2019). Testing all these new wells in the laboratory is unrealistic. Using simple statistical analyses, we have shown that even with its limited accuracy, the mean exposure post switching based on kit measurements is not much higher than exposure post switching based on laboratory measurements. If a slight increase in arsenic exposure of a small proportion of the population (around 2%) is acceptable, then kits provide a cheap alternative of reducing arsenic exposure for the overall population. Widespread well switching could significantly reduce arsenic exposure in Bangladesh in the short term and until more sustainable solutions are developed.

6. References

Ahmad, A., Bhattacharya, P., 2019. Arsenic in Drinking Water: Is 10 µg/L a Safe Limit? Curr. Pollut. Reports. https://doi.org/10.1007/s40726-019-0102-7

Ahmad, A., van der Wens, P., Baken, K., de Waal, L., Bhattacharya, P., Stuyfzand, P., 2020. Arsenic reduction to <1 µg/L in Dutch drinking water. Environ. Int. https://doi.org/10.1016/j.envint.2019.105253

Ahmed, M.F., Ahuja, S., Alauddin, M., Hug, S.J., Lloyd, J.R., Pfaff, A., Pichler, T., Saltikov, C., Stute, M., Van Geen, A., 2006. Ensuring safe drinking water in Bangladesh. Science (80-.). https://doi.org/10.1126/science.1133146

Ahsan, H., Chen, Y., Parvez, F., Argos, M., Hussain, A.I., Momotaj, H., Levy, D., Van Geen, A., Howe, G., Graziano, J., 2006. Health Effects of Arsenic Longitudinal Study (HEALS): Description of a multidisciplinary epidemiologic investigation. J. Expo. Sci. Environ. Epidemiol. 16, 191–205. https://doi.org/10.1038/sj.jea.7500449

Barnwal, P., van Geen, A., von der Goltz, J., Singh, C.K., 2017. Demand for environmental quality information and household response: Evidence from well-water arsenic testing. J. Environ. Econ. Manage. 86, 160–192. https://doi.org/10.1016/j.jeem.2017.08.002

Bennear, L., Tarozzi, A., Pfaff, A., Balasubramanya, S., Matin Ahmed, K., van Geen, A., 2013. Impact of a randomized controlled trial in arsenic risk communication on household water-source choices in Bangladesh. J. Environ. Econ. Manage. 65, 225–240. https://doi.org/10.1016/j.jeem.2012.07.006

BenYishay, A., Mobarak, A.M., 2019. Social Learning and Incentives for Experimentation and Communication. Rev. Econ. Stud. 86, 976–1009. https://doi.org/10.1093/restud/rdy039

Chen, Y., van Geen, A., Graziano, J.H., Pfaff, A., Madajewicz, M., Parvez, F., Hussain, A.Z.M.I., Slavkovich, V., Islam, T., Ahsan, H., 2007. Reduction in Urinary Arsenic Levels in Response to Arsenic Mitigation Efforts in Araihazar, Bangladesh. Environ. Health Perspect. 115, 917–923. https://doi.org/10.1289/ehp.9833

Cheng, Z., Zheng, Y., Mortlock, R., Van Geen, A., 2004. Rapid multi-element analysis of groundwater by high-resolution inductively coupled plasma mass spectrometry. Anal. Bioanal. Chem. 379, 512–518. https://doi.org/10.1007/s00216-004-2618-x

Dey, N.C., Saha, R., Parvez, M., Bala, S.K., Islam, A.S., Paul, J.K., Hossain, M., 2017. Sustainability of groundwater use for irrigation of dry-season crops in northwest Bangladesh. Groundw. Sustain. Dev. 4, 66–77. https://doi.org/10.1016/j.gsd.2017.02.001

Fendorf, S., Michael, H.A., Van Geen, A., 2010. Spatial and temporal variations of groundwater arsenic in South and Southeast Asia. Science (80-.). https://doi.org/10.1126/science.1172974

Flanagan, S. V., Johnston, R.B., Zheng, Y., 2012. Aarsenic dans l'eau des puits tubulaires au Bangladesh: Impacts sanitaires et économiques, et implications en vue de sa réduction. Bull. World Health Organ. 90, 839–846. https://doi.org/10.2471/BLT.11.101253

Gelman, A., Trevisani, M., Lu, H., Van Geen, A., 2004. Direct Data Manipulation for Local Decision Analysis as Applied to the Problem of Arsenic in Drinking Water from Tube Wells in Bangladesh. Risk Anal. 24, 1597–1612. https://doi.org/10.1111/j.0272-4332.2004.00553.x

George, C.M., Graziano, J.H., Mey, J.L., van Geen, A., 2012. Impact on arsenic exposure of a growing proportion of untested wells in Bangladesh. Environ. Heal. 11. https://doi.org/10.1186/1476-069x-

11-7

Huhmann, B.L., Harvey, C.F., Navas-Acien, A., Graziano, J., Parvez, F., Chen, Y., Argos, M., Ahmed, A., Hasan, A.K.M.R., Ahsan, H., van Geen, A., 2019. Changes in arsenic exposure in Araihazar, Bangladesh from 2001 through 2015 following a blanket well testing and education campaign. Environ. Int. 125, 82–89. https://doi.org/10.1016/j.envint.2019.01.026

Inauen, J., Tobias, R., Mosler, H.J., 2014. The role of commitment strength in enhancing safe water consumption: Mediation analysis of a cluster-randomized trial. Br. J. Health Psychol. 19, 701–719. https://doi.org/10.1111/bjhp.12068

Jakariya, M., Vahter, M., Rahman, M., Wahed, M.A., Hore, S.K., Bhattacharya, P., Jacks, G., Persson, L.Å., 2007. Screening of arsenic in tubewell water with field test kits: Evaluation of the method from public health perspective. Sci. Total Environ. 379, 167–175. https://doi.org/10.1016/j.scitotenv.2006.11.053

Jamil, N.B., Feng, H., Ahmed, K.M., Choudhury, I., Barnwal, P., Van Geen, A., 2019. Effectiveness of Different Approaches to Arsenic Mitigation over 18 Years in Araihazar, Bangladesh: Implications for National Policy. Environ. Sci. Technol. 53, 5596–5604. https://doi.org/10.1021/acs.est.9b01375

Krupoff, M., Mobarak, A.M., Van Geen, A., 2020. Evaluating strategies to reduce arsenic poisoning in south asia: A view from the social sciences. Asian Dev. Rev. 37, 21–44. https://doi.org/10.1162/adev_a_00148

Madajewicz, M., Pfaff, A., van Geen, A., Graziano, J., Hussein, I., Momotaj, H., Sylvi, R., Ahsan, H., 2007. Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh. J. Dev. Econ. 84, 731–754. https://doi.org/10.1016/j.jdeveco.2006.12.002

Mihajlov, I., Stute, M., Schlosser, P., Mailloux, B.J., Zheng, Y., Choudhury, I., Ahmed, K.M., van Geen, A., 2016. Recharge of low-arsenic aquifers tapped by community wells in Araihazar, Bangladesh, inferred from environmental isotopes. Water Resour. Res. 52, 3324–3349. https://doi.org/10.1002/2015WR018224

Mozumder, M.R.H., Michael, H.A., Mihajlov, I., Khan, M.R., Knappett, P.S.K., Bostick, B.C., Mailloux, B.J., Ahmed, K.M., Choudhury, I., Koffman, T., Ellis, T., Whaley-Martin, K., San Pedro, R., Slater, G., Stute, M., Schlosser, P., Geen, A., 2020. Origin of Groundwater Arsenic in a Rural Pleistocene Aquifer in Bangladesh Depressurized by Distal Municipal Pumping. Water Resour. Res. 56. https://doi.org/10.1029/2020WR027178

Pattanayak, S.K., Pfaff, A., 2009. Behavior, Environment, and Health in Developing Countries: Evaluation and Valuation. Annu. Rev. Resour. Econ. 1, 183–217. https://doi.org/10.1146/annurev.resource.050708.144053

Quansah, R., Armah, F.A., Essumang, D.K., Luginaah, I., Clarke, E., Marfoh, K., Cobbina, S.J., Nketiah-Amponsah, E., Namujju, P.B., Obiri, S., Dzodzomenyo, M., 2015. Association of Arsenic with Adverse Pregnancy Outcomes/Infant Mortality: A Systematic Review and Meta-Analysis. Environ. Health Perspect. 123, 412–421. https://doi.org/10.1289/ehp.1307894

Reddy, R.R., Rodriguez, G.D., Webster, T.M., Abedin, M.J., Karim, M.R., Raskin, L., Hayes, K.F., 2020. Evaluation of arsenic field test kits for drinking water: Recommendations for improvement and implications for arsenic affected regions such as Bangladesh. Water Res. 170, 115325. https://doi.org/10.1016/j.watres.2019.115325

Smith, A.H., Smith, M.M.H., 2004. Arsenic drinking water regulations in developing countries with extensive exposure, in: Toxicology. Elsevier, pp. 39–44. https://doi.org/10.1016/j.tox.2004.02.024

Tarozzi, A., Maertens, R., Ahmed, K.M., van Geen, A., 2020. Demand for Information on Environmental Health Risk, Mode of Delivery, and Behavioral Change: Evidence from Sonargaon, Bangladesh, Demand for Information on Environmental Health Risk, Mode of Delivery, and Behavioral Change: Evidence from Sonargaon, Bangladesh. World Bank, Washington, DC. https://doi.org/10.1596/1813-9450-9194

Van Geen, A., Ahmed, E.B., Pitcher, L., Mey, J.L., Ahsan, H., Graziano, J.H., Ahmed, K.M., 2014. Comparison of two blanket surveys of arsenic in tubewells conducted 12years apart in a 25km2 area of Bangladesh. Sci. Total Environ. 488–489, 484–492. https://doi.org/10.1016/j.scitotenv.2013.12.049

Van Geen, A., Ahsan, H., Horneman, A.H., Dhar, R.K., Zheng, Y., Hussain, I., Ahmed, K.M., Gelman, A., Stute, M., James Simpson, H., Wallace, S., Small, C., Parvez, F., Slavkovich, V., Loiacono, N.J., Becker, M., Cheng, Z., Momotaj, H., Shahnewaz, M., Seddique, A.A., Graziano, J.H., 2002. Promotion of well-switching to mitigate the current arsenic crisis in Bangladesh. Bull. World Health Organ. 80, 732–737. https://doi.org/10.1590/S0042-96862002000900010

Van Geen, A., Cheng, Z., Seddique, A.A., Hoque, M.A., Gelman, A., Graziano, J.H., Ahsan, H., Parvez, F., Ahmed, K.M., 2005. Reliability of a commercial kit to test groundwater for arsenic in Bangladesh. Environ. Sci. Technol. 39, 299–303. https://doi.org/10.1021/es0491073

van Geen, A., Zheng, Y., Versteeg, R., Stute, M., Horneman, A., Dhar, R., Steckler, M., Gelman, A., Small, C., Ahsan, H., Graziano, J.H., Hussain, I., Ahmed, K.M., 2003. Spatial variability of arsenic in 6000 tube wells in a 25 km ^{2} area of Bangladesh. Water Resour. Res. 39. https://doi.org/10.1029/2002WR001617

Yu, W.H., Harvey, C.M., Harvey, C.F., 2003. Arsenic in groundwater in Bangladesh: A geostatistical and epidemiological framework for evaluating health effects and potential remedies. Water Resour. Res. 39. https://doi.org/10.1029/2002WR001327

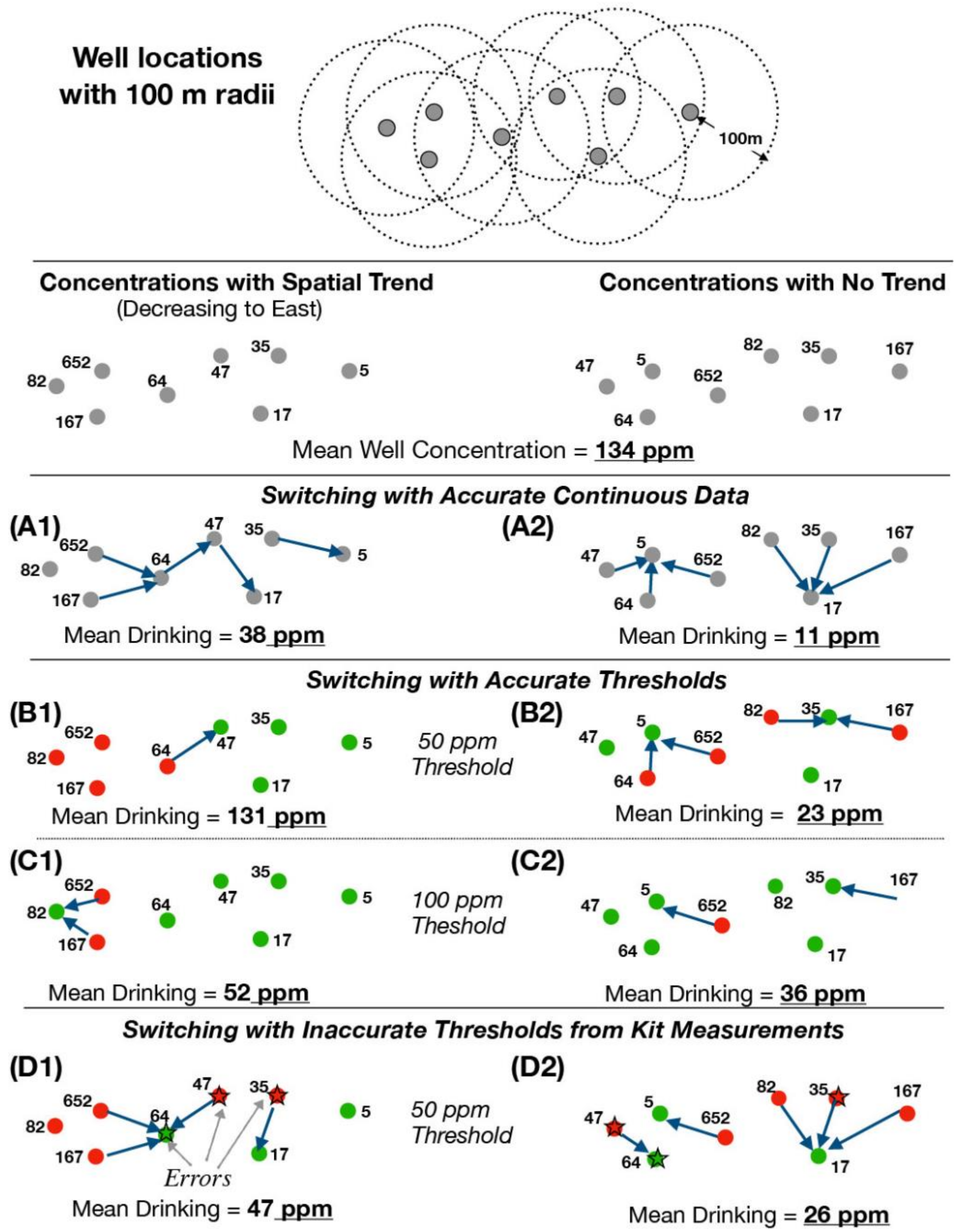

Figure 1: Schematic of a hypothetical group of wells and their arsenic concentrations that illustrates well switching patterns. **Top panel**: Positions and arsenic concentrations for 8 hypothetical wells (shown in gray circle) and the 100 m radii in which switching is possible. **First column:** Examples were arsenic concentrations are correlated space, generally decreasing to the east. **Second column:** Examples were arsenic concentrations are uncorrelated in space. **Row A:** The ideal base case where switches are based on accurate continuous arsenic measurements. **Row B and C:** Switches are based on thresholds. **Row D:** Switches are based on kit measurements that mis-assign some wells to the wrong category.

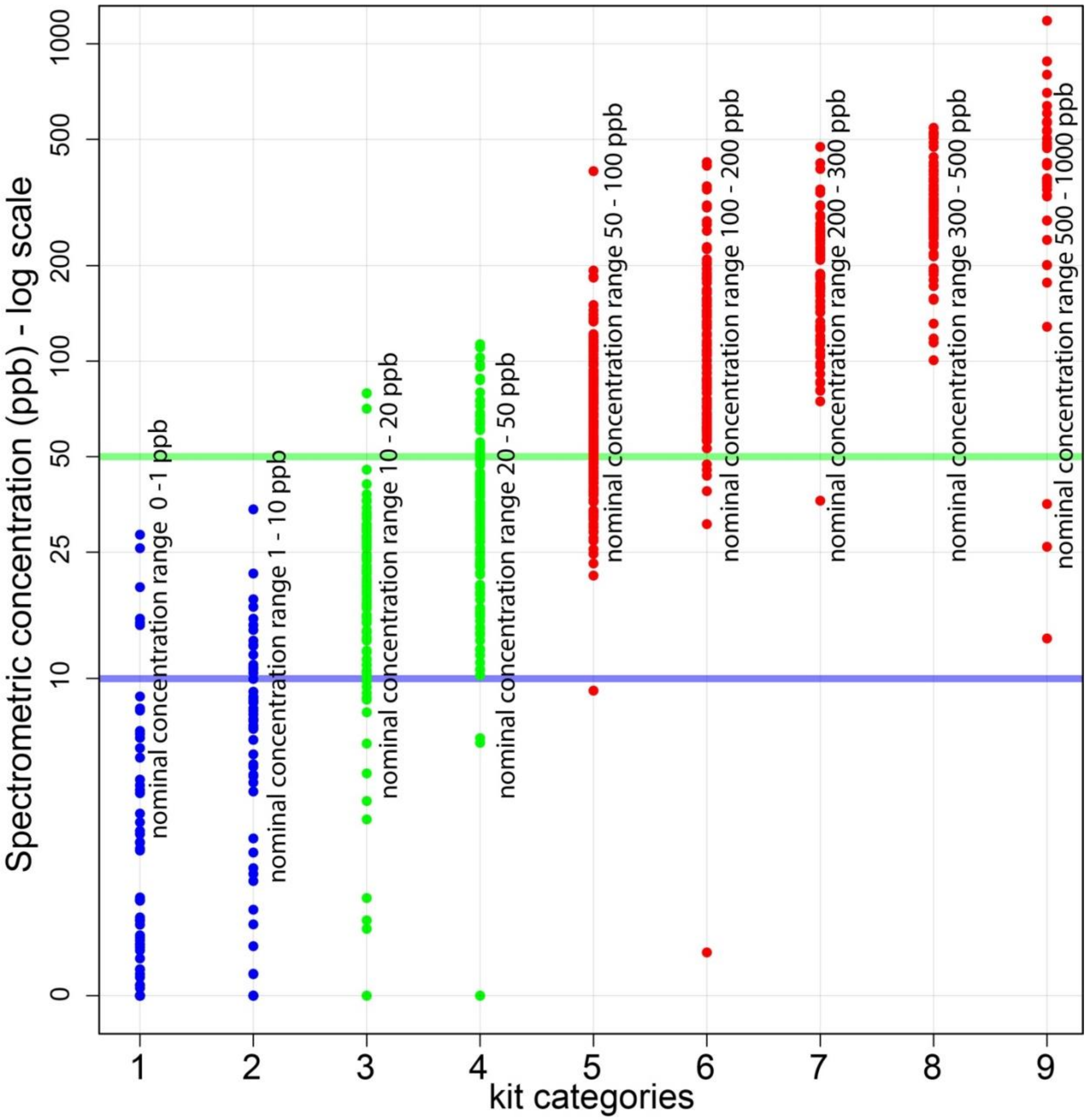

Figure 2: Distribution of arsenic concentrations (y-axis) within kit categories (x-axis) for the 900 wells with paired spectrometric and kit measurements. Kit categories shown in blue and green are classified as uncontaminated and kit categories shown in red are classified as contaminated.

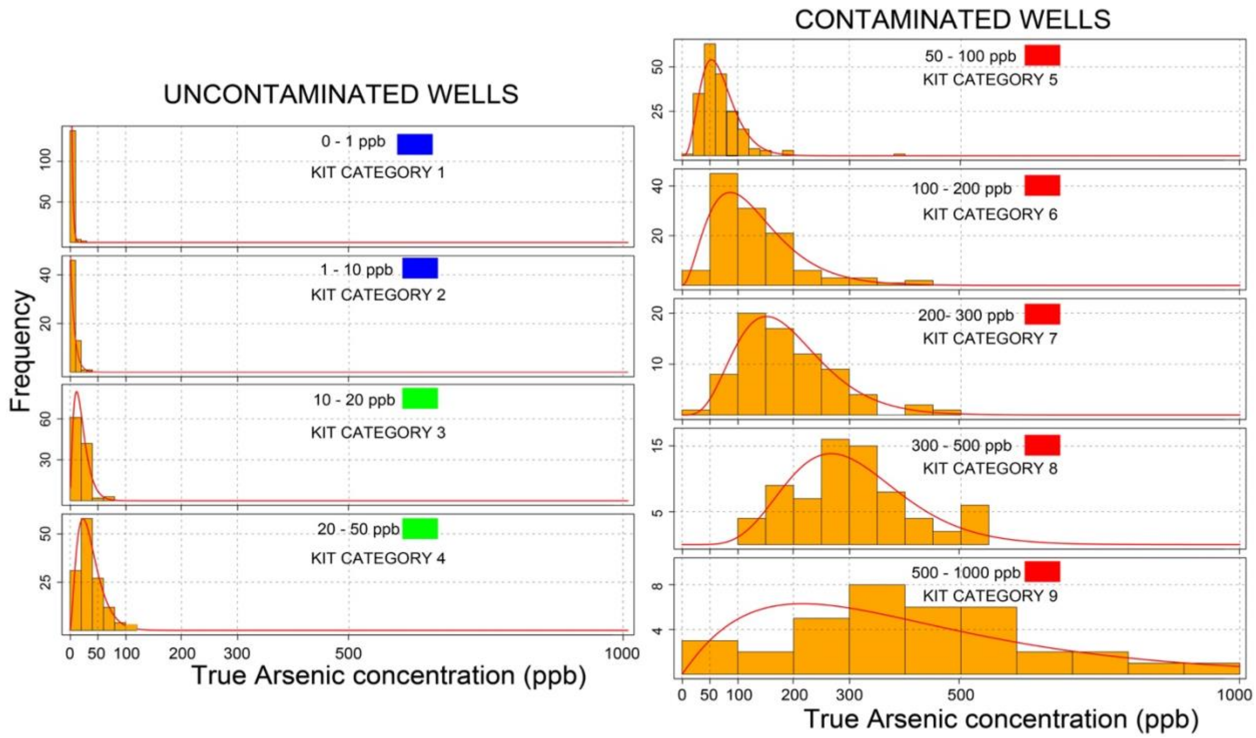

Figure 3: Frequency distribution (orange bars) of the arsenic concentration (x-axis) for the respective nominal kit categories. The red line is the best fit parametric distribution to the data for each kit categories. The placards posted on well based on the kit categories are also shown in each panel. Wells with blue and green placards are considered as uncontaminated wells and wells with red placards are considered as contaminated wells.

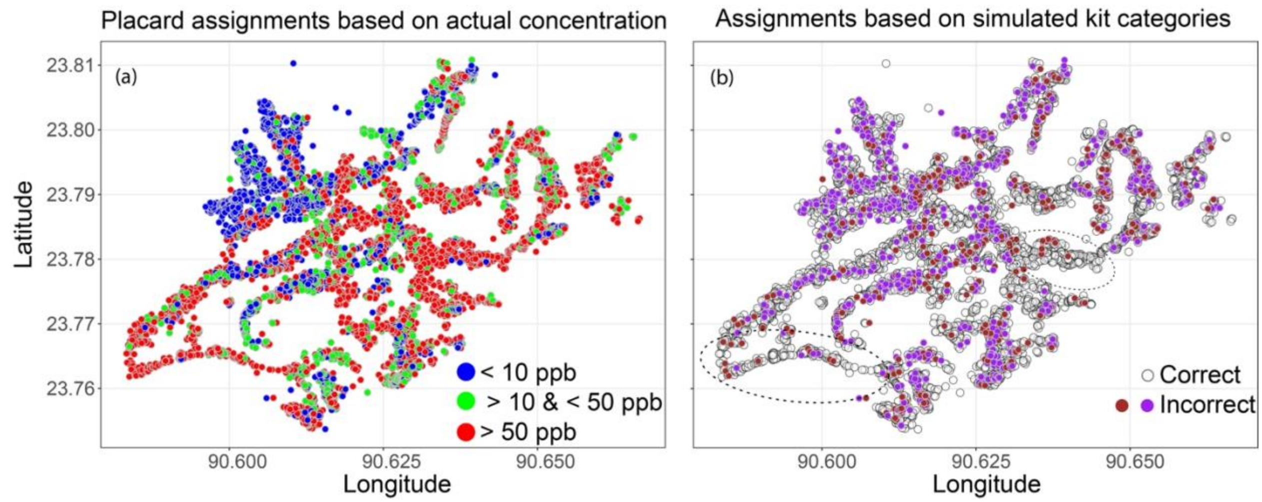

Figure 4: (a) Arsenic concertation of 6595 Araihazar wells measured by spectrometric method. (b) Correct (white circle with black border), and incorrect assignments (purple and brown circles) for simulated categorization based on kit measurements (see table 2). Wells with $As < 10$ ug/L that are labeled as green or red and wells with $10 < As < 50$ that are labeled as red are shown in purple. Wells with $10 < As < 50$ that are labeled as blue and wells with $As > 50$ labeled as blue and green are shown in brown. Two regions with large proportion of correct assignments are highlighted in black ellipses.

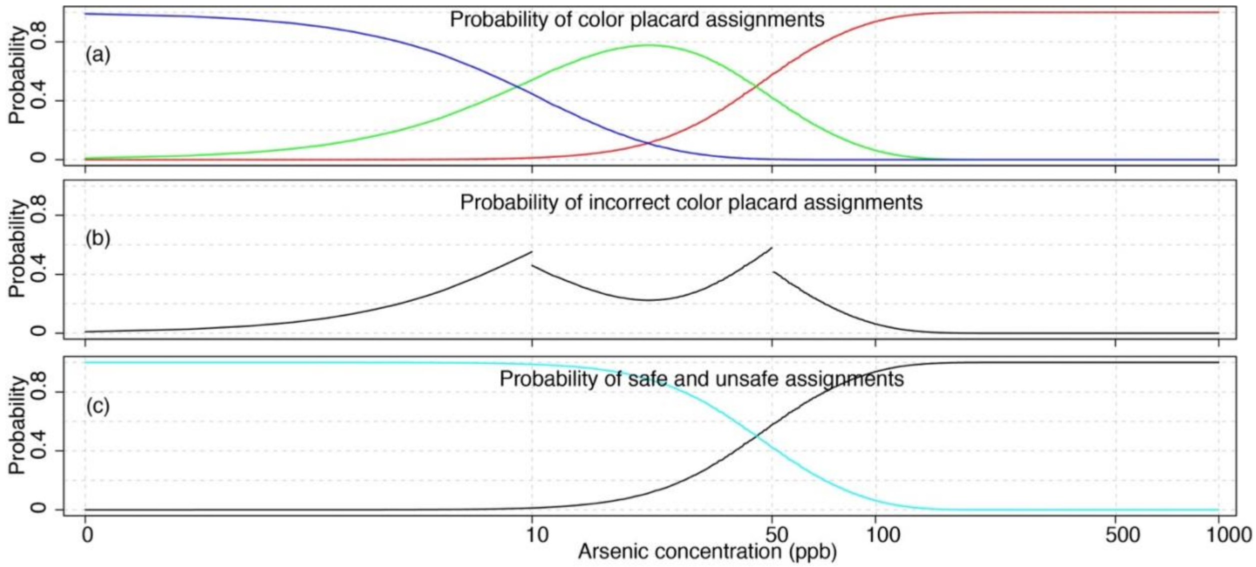

Figure 5: (a) Probability of assigning the different color placards from kit measurements of arsenic as a function of arsenic concentration. (b) Probability of assigning incorrect color placard as a function of arsenic concentration. For wells with As < 10 ug/L , probability of incorrect assignment is defined as the sum of assigning green and red placards. For well with As > 10 ug/L and As < 50 ug/L , probability of incorrect assignment is defined as the sum of assigning blue and red placards. For wells with As > 50 ug/L , probability of incorrect assignment is defined as the sum of assigning blue and green placards. Please refer to table 3 for mean probability of different color assignments for 0-10 ug/L , 10-50 ug/L and >50 ug/L . (c) Probability of assigning a well as safe (light blue) and unsafe (black) as a function of arsenic concentration.

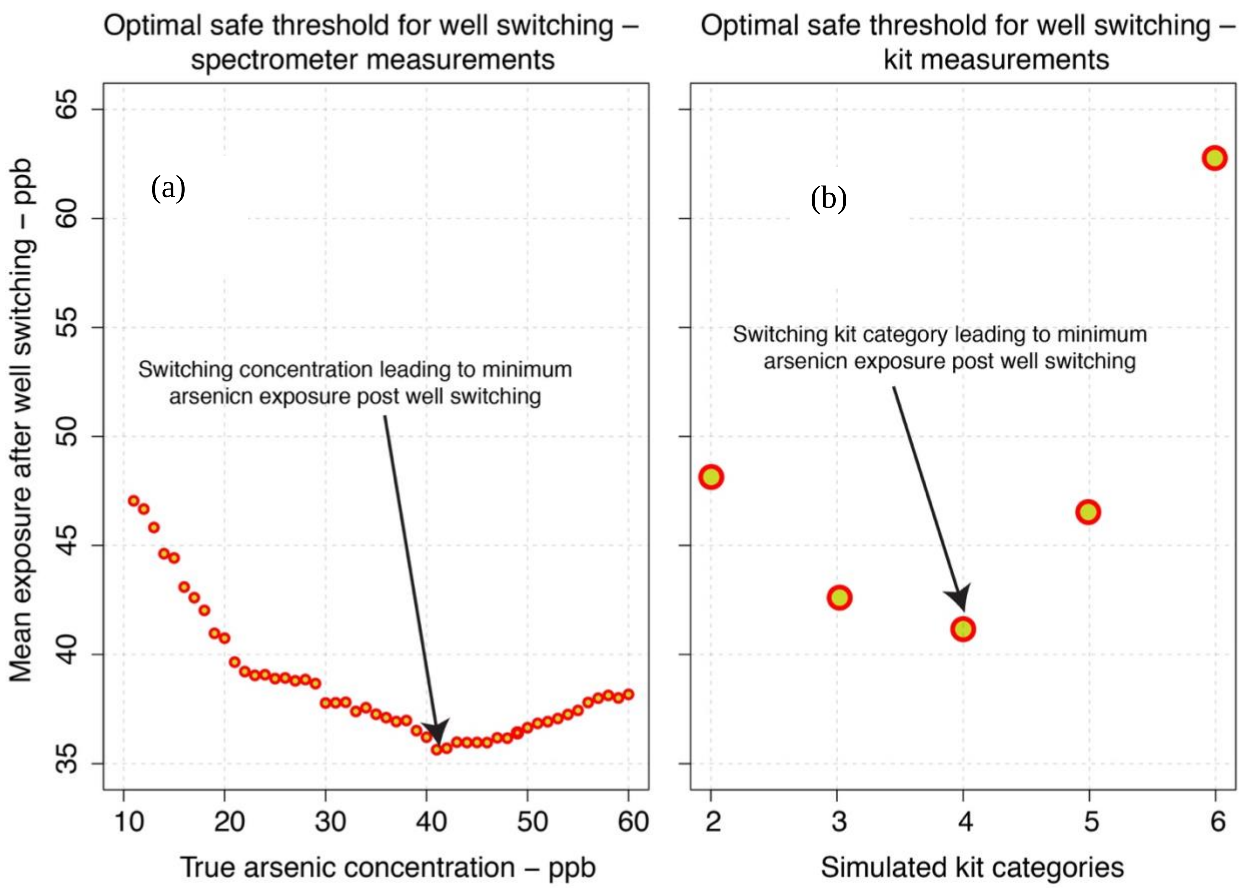

Figure 6: (a) Mean exposure post switching for different "safe" thresholds. The minimum exposure is at 40 ug/L . (b) Mean exposure post switching based on categorical kit measurements. The minimum exposure post switching is observed if wells with categories 1-4 are labeled green and wells with categories 5-9 are labeled red.

Table 1: Density and conditional probabilities of the nominal kit categories for the laboratory (spectrometric measured) concentration of 100 ug/L. Please refer to Figure 3 for visual reference of the density. Kit categories (5, 6 and 7) with high probability are highlighted in bold. The conditional probabilities are calculated using equation 1.

Nominal Kit categories	**Density**	**Conditional Probabilities**
Kit category 1 (nominal range 0 - 1ug/L)	<0.0001	P(Kit category = 1 | As = 100 ug/L) <0.001
Kit category 2 (nominal range 1 - 10ug/L)	<0.0001	P(Kit category = 2 | As = 100 ug/L) <0.001
Kit category 3 (nominal range 10 - 20ug/L)	<0.0001	P(Kit category = 3 | As = 100 ug/L) <0.001
Kit category 4 (nominal range 20 - 50ug/L)	<0.0001	P(Kit category = 4 | As = 100 ug/L) <0.001
Kit category 5 (nominal range 50 - 100ug/L)	0.005	**P(Kit category = 5 | As = 100 ug/L) = 0.3**
Kit category 6 (nominal range 100 - 200ug/L)	0.008	**P(Kit category = 6 | As = 100 ug/L) = 0.45**
Kit category 7 (nominal range 200 - 300ug/L)	0.003	**P(Kit category = 7 | As = 100 ug/L) = 0.2**
Kit category 8 (nominal range 300 - 500ug/L)	0.0003	P(Kit category = 8 | As = 100 ug/L) = 0.04
Kit category 9 (nominal range 500 - 1000ug/L)	<0.0001	P(Kit category = 9 | As = 100 ug/L) < 0.0001

Table 2: Possible switching scenarios based on the probability of correct, false positive and false negative nominal kit category assignments. *i* is the different kit category and θ is the true arsenic concentration. Note: Actual switching only takes place when the well to be switched to lies in a 100 m radius of the well that is being switched from.

Type of switching	**Probability of switching**	**Description**
Ideal Switching	$P(i=\{5,6,7,8,9\} \mid \theta>50)*P(i=\{1,2,3,4\} \mid \theta<50)$	Switched from correctly identified unsafe well to correctly identified safe well
Bad Switching	$P(i=\{5,6,7,8,9\} \mid \theta>50)*P(i=\{1,2,3,4\} \mid \theta>50)$	Switched from correctly identified unsafe well to incorrectly identified safe well (false negative)
Very bad switching	$P(i=\{5,6,7,8,9\} \mid \theta<50)*P(i=\{1,2,3,4\} \mid \theta>50)$	Switched from incorrectly identified safe well (false negative) to incorrectly identified unsafe well (false negative)
Unnecessary switching	$P(i=\{5,6,7,8,9\} \mid \theta<50)*P(i=\{1,2,3,4\} \mid \theta<50)$	Switched from incorrectly identified safe well (false positive) to correctly identified safe well
Missed switching	$P(i=\{1,2,3,4\} \mid \theta>50)$	False negative identification of unsafe wells. Therefore, no switching

Table 3: Summary statistics of exposure post switching and percentage of consumers experiencing change in arsenic exposure for the different switching scenarios. Before switching the exposure is the average arsenic concentration across wells 134 ug/L. Scenario B2 and B4 have same values for the different rows as the optimal threshold (B4) for switching between uncontaminated and contaminated (red and green) wells is observed when wells with kit categories 1-4 are labeled green and categories 5-9 are labeled red (see Figure 6).

	SPECTROMETRIC MEASUREMENTS					**SIMULATED KIT CATEGORIES**				**NO SPATIAL CORRELATION**	
	GROUP A					**GROUP B**				**GROUP C**	
	Switching with perfect continuous measurements	Switching with uncertainty in continuous measurements	Switching based on 2 categories of contamination	Switching based on 3 categories of contamination	Optimal switching threshold	Switching from higher to lowest kit categories	Switching based on 2 categories	Switching based on 3 categories	Optimal switching threshold	Spectrometr ic	kit
	Scenario A1	Scenario A2	Scenario A3	Scenario A4	Scenario A5	Scenario B1	Scenario B2	Scenario B3	Scenario B4	Scenario C1	Scenario C2
Mean exposure post switching**	17 ug/L	18 ug/L	37 ug/L	34 ug/L	35 ug/L	24 ug/L	41 ug/L	35 ug/L	41 ug/L	6 ug/L	8 ug/L
Arsenic exposure decreased	84%	83%	43%	43%	43%	69%	42%	44%	42%	84%	76%
Arsenic exposure did not change	16%	17%	57%	57%	57%	25%	55%	54%	55%	16%	21%
Arsenic exposure increased	0%	0%	0%	0%	0%	6%	4%	2%	4%	0%	3%
Percent of post-switched well that itself switches to another well	17%	17%	0%	0%	0%	10%	0%	0%	0%	10%	7.5%