ENVIRONMENTAL HEALTH PERSPECTIVES

http://www.ehponline.org

Association of Supply Type with Fecal Contamination of Source Water and Household Stored Drinking Water in Developing Countries: A Bivariate Meta-analysis

Katherine F. Shields, Robert E.S. Bain, Ryan Cronk, Jim A. Wright, and Jamie Bartram

http://dx.doi.org/10.1289/ehp.1409002

Received: 25 July 2014
Accepted: 6 May 2015
Advance Publication: 8 May 2015

Environ Health Perspect doi: 10.1289/ehp.1409002
Advance Publication: Not Copyedited

Association of Supply Type with Fecal Contamination of Source Water and Household Stored Drinking Water in Developing Countries: A Bivariate Meta-analysis

Katherine F. Shields,[1] Robert E.S. Bain,[1,2] Ryan Cronk,[1] Jim A. Wright,[3] and Jamie Bartram[1]

[1]The Water Institute, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, USA; [2]UNICEF, New York, New York, USA; [3]Geography and Environment, University of Southampton, Southampton, United Kingdom

Address correspondence to Katherine F. Shields, The Water Institute, University of North Carolina at Chapel Hill, 148 Rosenau Hall, CB #7431, Chapel Hill, NC 27599-7431 USA. Telephone: (734) 649-3295. E-mail: kshields@email.unc.edu

Running title: Source and stored water contamination

Acknowledgments: We thank the Howard W. Odum Institute for Research in Social Science at the University of North Carolina, Chapel Hill for help with the statistical methods. Thanks to Mike Fisher and Kristen Downs for their helpful comments on drafts of this paper. This work was supported by the Water Working Group of the WHO/UNICEF JMP (www.wssinfo.org) and by WHO (www.who.int). The funders had no role in study design, data collection and analysis, or decision to publish.

Disclaimer: The authors alone are responsible for the views expressed in this publication and they do not necessarily represent the decisions or policies of UNICEF, the World Health Organization or UNC.

Competing financial interests: JB is a member of the Water Working Group and is an unpaid advisor to both WHO and UNICEF. The other authors declare they have no actual or potential competing financial interests.

Abstract

Background: Access to safe drinking water is essential for health. Monitoring access to drinking water focuses on water supply type at the source, but there is limited evidence on whether quality differences at the source persist in water stored in the household.

Objectives: To assess the extent of fecal contamination at the source and in household stored water (HSW) and explore the relationship between contamination at each of these sampling points and water supply type.

Methods: A bivariate random-effects meta-analysis of 45 studies, identified through a systematic review, that reported either the proportion of samples free of fecal indicator bacteria and/or individual sample bacteria counts for source and HSW, disaggregated by supply type.

Results: Water quality deteriorated substantially between source and stored water. Mean percentage of contaminated samples (noncompliance) at the source was 46% (95% CI: 33, 60%) while mean noncompliance in HSW was 75% (95% CI: 64, 84%). Water supply type was significantly associated with noncompliance at the source ($p < .001$) and in HSW ($p = 0.03$). Source water (OR = 0.2; 95% CI: 0.1, 0.5) and HSW (OR = 0.3; 95% CI: 0.2, 0.8) from piped supplies had significantly lower odds of contamination when compared to non-piped water, potentially due to residual chlorine.

Conclusions: Piped water is less likely to be contaminated compared to other water supply types at both the source and in HSW. A focus on upgrading water services to piped supplies may help improve safety, including for those drinking stored water.

Introduction

The health consequences of drinking fecally contaminated water, particularly for young children and immunocompromised individuals, have long been recognized (Gerba et al. 1996). International development initiatives, including the International Drinking-water Supply and Sanitation decade in the 1980s and the more recent Millennium Development Goals, have focused global policy attention on access to safe water (Bradley and Bartram 2013). Access to safe drinking water is monitored by the Joint Monitoring Programme for Water Supply and Sanitation (JMP) of WHO and UNICEF using the dichotomous indicator of the proportion of the population using an "improved" drinking water supply, which includes piped water, boreholes, protected springs and dug wells, and rainwater. Unprotected springs and dug wells, carts with small tanks, tanker trucks, and surface water are considered "unimproved" (Bartram et al. 2014). Although WHO and UNICEF declared that the world had met the drinking water target in 2010, as assessed by use of this indicator, they cautioned that it is likely that the number of people using *safe* water had been over-estimated (WHO/UNICEF 2012). Assessments which take source water quality into account suggest that between 1.8 and 3 billion people, or 28%-47% of the global population, used unsafe water or water from sanitarily unsafe supplies in 2010 (Onda et al. 2012). Because many improved supplies are remote from households, they require transportation and storage of drinking water. Even when water is piped into a dwelling or yard, water storage may be required due to intermittent or unreliable supply.

Recognition of contamination during transport and household storage has sparked debate on the relative importance of water quality and treatment at the source versus point of use (PoU), the point at which drinking water is consumed (Boisson et al. 2013; Clasen and Cairncross 2004;

Mintz et al. 2001; Trevett et al. 2004; VanDerslice and Briscoe 1993). Some have probed the health significance of intra- versus extra-household contamination. For example, VanDerslice and Briscoe (1993) suggest most household members have immunity to pathogens already circulating in the household. Extra-household contamination poses a greater health risk, they argued, as it has the potential of bringing new pathogens into the household. Even if no immunity were present, transmission of pathogens via stored water would be inefficient relative to other transmission pathways. In contrast, Trevett et al. (2005) constructed a model for contamination of water stored in the household, theorizing that VanDerslice and Briscoe obtained their results because their study population had a low risk for contamination of stored water. Highly contaminated stored water would have a greater effect on health they argued. Others have explored the treatment options, suggesting that household treatment is essential, at least for the present, as source water treatment is more time and resource intensive (Mintz et al. 2001). Finally, Clasen and Bastable (2003) and Wright et al. (2004) have emphasized the need to measure and monitor quality at the PoU in addition to the source.

In their review, Wright et al. (2004) found that, in half of the studies included, contamination was greater at the PoU than at the source, and in no case did microbiological water quality improve significantly from the source to the PoU. Wright et al. (2004) were unable to explore differences between types of water supplies; in their analysis unprotected well water and household connections delivering chlorinated water were treated as equal. Thus, our goal here is to build upon the work of Wright et al. (2004) by exploring contamination at the water source separately from contamination of household stored water (HSW), and whether factors such as water provenance, rural/urban setting, and indicator organism have a differential effect on contamination at each location. In light of the findings of Bain et al. (2014b) that water supply type is strongly associated with noncompliance (i.e. the percentage of water samples contaminated by fecal indicator bacteria (FIB)), we describe the supply types in as much detail as possible instead of aggregating them into improved and unimproved categories only. We use 'water source' to mean the point of collection or receipt by the household. This includes water as taken from a river, a handpump, community tap stand, or tap in a household as well as water as it is received from a vendor or tanker truck. Hereafter, we use the term HSW, referring to any water stored in the home, rather than water at the PoU, referring to the point that water is drawn from immediately prior to consumption or use. The distinction is important because water may be transferred to another container and it is unclear how long water is stored prior to consumption by household members.

Methods

Study Selection

Studies used in this review and meta-analysis are a subset of those included in the systematic review by Bain et al. (2014b) who analyzed 319 studies based on the inclusion criteria that: studies reported on water quality at the source or in HSW; would not be classified as surface water by the JMP; and included extractable data on thermotolerant coliforms (TTC) or *E. coli,* with sample volumes of no less than 10 mL and at least 10 samples from different water supplies of a given type. Articles in English, Spanish, French and Portuguese, describing studies in developing countries, as defined by the MDG regions, and published between January 1990 and August 2013 were included.

Abstracts of the 319 articles analyzed by Bain et al. (2014b) were screened for those reporting data on water quality at both the source *and* in HSW. The most common definition of HSW was water sampled from a household storage container. Some studies provided additional detail of container type (household water jug, opaque containers, bucket, etc.) and storage duration (for example no more than 24 hours). A few studies referred simply to household water, with storage implied in the sampling description. Where abstracts were unclear, methods sections were reviewed. Full texts were reviewed for articles which reported sampling both water sources and HSW. If methods sections indicated that water quality data at both the source and in HSW were collected, but either data were not reported from both sampling points, or reported data were not disaggregated by water supply type, authors were contacted at least twice to request data. Studies were excluded if they did not report data about both source and stored water, if the data were not disaggregated by supply type, or if fewer than 10 samples were taken at either the source or from HSW. Rainwater collection, which acts as both a source and HSW, was excluded.

Some studies sampled pairs, taking one water sample from stored water in a household and one sample from the water source used by the household. Other studies sampled a number of HSW and all available sources, creating pairs *post hoc* by linking stored water data of a particular household to source water data based on reported supply type.

Study information extracted by Bain et al. (2014b) and used in this analysis included: (i) supply type (unprotected well, unprotected spring, unspecified well or spring, protected well, protected spring, borehole, piped, surface, tanker truck, or bottled water); (ii) source treatment (i.e. reported chlorination and assessed residual chlorine) and household water treatment (HWT) (i.e. boiling, filtration, chlorination); (iii) compliance (percent of samples 10mL or larger free of *E. coli* or thermotolerant coliforms (TTC)), (iv) mean, geometric mean and/or median level of contamination by water supply type; (v) number of samples collected at source and from stored water; (vi) location (urban or rural as defined by the authors, ''mixed' if the setting was mixed); (vii) study country; (viii) year of publication; (ix) study design; and (x) study quality information as described in Bain et al. (2014b) (see Supplemental Material, Table S1). Studies were assigned a point for each item in Table S1 resulting in a study quality score of between 0 and 13, and then broken into terciles of "low"(<7), "medium" (7-9) and "high" (>9) quality. World Bank classifications from 2013 were used to determine country income levels (World Bank 2013). For intervention studies, baseline water quality data for both the intervention and control groups were used where possible. If baseline data were not available and for case-control studies, water quality data from the control group were used.

Meta-analysis

Meta-analysis was used to explore factors which were associated with differences in water quality including water supply characteristics, study setting, study characteristics and reporting format. We used bivariate random effects meta-analysis and meta-regression to analyze noncompliance at the source, noncompliance of HSW, and the odds ratio between the two simultaneously. Studies where water was sampled from the drinking cup rather than from household storage, studies in which all sampled HSW was known to have been treated in the household, and studies which reported only central tendency of FIB were excluded from the meta-analysis.

Analysis was completed using PROC GLIMMIX in SAS 9.4 (SAS Institute). Number of "events" (water samples contaminated at a given sampling point) compared to "trials" (total number of samples analyzed at a given sampling point) were modeled using a bivariate distribution and a logit link function. Degrees of freedom were set at 1000 in order to produce z-statistics rather than a student-t statistic (Reitsma et al. 2005). Covariates were analyzed as interactions with each sampling point (source and HSW) in order to avoid assuming a covariate had the same effect at both.

In order to avoid assuming a fixed relationship between quality at the source and quality of HSW for any given water supply, the water sampling point (source or HSW) was used to model random effects. Pooled logit estimates at both sampling points were back transformed to report mean noncompliance on the scale of the original data. Overall significance levels for each covariate were calculated using Type III tests of the fixed effects of the interaction between sampling point (source or household stored) and the covariate of interest. Odds ratios were calculated to compare levels of categorical variables (e.g. improved versus unimproved water supplies) at both the source and in HSW and to compare noncompliance between source and HSW, after adjustment for water supply type. The 95% confidence regions for mean noncompliance were calculated using the geometric relationships between variance and correlation (Pakula, L 2008) and plotted using SAS 9.4

We used the multivariate I_R^2 statistic developed by Jackson et al. (2012) (as described in Zhou and Dendukuri 2014) to quantify study heterogeneity. To examine publication bias, log odds ratios were calculated between HSW and source, with log odds of greater than zero indicating higher noncompliance in HSW. A funnel plot was created and the trim and fill method applied using STATA 12 (StataCorp).

Results

Study Characteristics

A total of 319 abstracts were screened and 114 full text articles assessed (Figure 1). Characteristics of the 45 included studies are summarized in Table 1. Descriptive statistics for included studies are presented in Table 2. Most studies were cross sectional (n = 27, 60%) while a quarter were intervention studies (n = 12, 27%). Seven of the MDG regions were represented in this review (see Supplemental Material, Figure S1) with most studies taking place in sub-Saharan Africa (n=24, 53%), 10 in South Asia (22%) and seven in Latin America and the Caribbean (16%). Low and lower-middle income countries dominated (n=21 and n=17 respectively). A similar number of studies had rural and urban settings (n = 20 and n = 21 respectively), with four studies having mixed settings. The majority of studies used *E. coli* (n=25, 56%) rather than TTC as the fecal indicator bacteria.

The 45 included studies reported on average 1.5 water supply types for a total of 65 water supply observations, with 10,934 total water samples taken at the source and 12,523 samples from HSW (Table 2). Thirty-two studies included source and HSW comparisons for only one supply type, while six studies included two supply types, six studies included three supply types and one study included four supply types. Nearly half of the water supply observations were piped (n=31, 49%) one third were 'other improved supplies' (n = 22) including boreholes, protected wells and protected springs. Six supplies were classified as unimproved (9%). In six cases, the protection status of wells could not be determined and these wells were treated as a separate supply category.

Of the 31 studies reporting on piped supplies, about half were reported to be chlorinated (n=14) while for an additional six supplies authors noted inconsistent or irregular chlorination (Table 2). Two supplies were not disinfected and for the remaining nine, chlorination status was not reported. All of the studies (n=6) reporting inconsistent or irregular chlorination assessed residual chlorine at the source, in HSW, or both. However, in only eight of the 14 supplies reported as chlorinated was residual chlorine detected. No other supply types were reported to be chlorinated. Residual chlorine was variously reported as an average or range of mg/L of chlorine or proportion of samples greater than 0.5 or 0.2 mg/L, preventing its inclusion in meta-analysis.

Two studies (Fiore et al. 2010 and Rosa et al. 2010) reported only on HSW treated in the home. Another study (Mertens et al. 1990) analyzed one boiled and one unboiled stored water sample from each household. We excluded these boiled water data because pairs of stored water samples from the same household are unlikely to be statistically independent. While many of the other studies mentioned household water treatment and storage practices, there was a lack of comparable data across the studies.

Between study analysis

Using a bivariate random effects model, water quality was found to be significantly worse in HSW compared to the source. Mean noncompliance at the source was found to be 46% (95% CI: 33, 60%) with mean noncompliance in HSW at 75% (95% CI: 64, 84%) (Table 3), an unadjusted odds ratio of 3.5 (2.5, 5.0) (Table 5). Noncompliance at the source ($p < .001$) and in HSW ($p = 0.03$) was found to be significantly associated with water supply type. At the source, mean noncompliance in piped water was 25% (95% CI: 15, 40%). In HSW, mean noncompliance for piped water was higher than at the source at 62% (95% CI: 44, 77%) (Table 3). Unprotected and unspecified wells had the highest mean noncompliance at both the source and in HSW. Protected and unprotected springs also had low rates of noncompliance at both the source and in the household, however the confidence intervals around these estimates were wide. After adjusting for supply type, HSW was found to have 2.3 higher odds of contamination than source water (95% CI: 1.4, 3.9) (Table 5).

Piped supplies had lower odds of being contaminated than other improved supplies (OR = 0.3; 95% CI: 0.1, 0.8) and all other supply types (OR = 0.2; 95% CI: 0.1, 0.5) at the source (Table 4). HSW from piped supplies had significantly lower odds of contamination when compared to all other supply types (OR = 0.3; 95% CI: 0.2, 0.8). Although the ellipses showing confidence limits for piped water versus water from other supplies overlap in Figure 2, the bivariate meta-regression indicated that odds ratios for noncompliant piped HSW versus source water was significantly different from the same odds ratio as calculated for other supply types

(Table 4). The confidence limits overlap due to the correlation between noncompliance at the source and in HSW.

Country income level (p = 0.03) was significantly associated with water quality at the source when all supply types are aggregated (Table 3). Lower-middle income countries had the highest mean noncompliance at 68% (95% CI: 47, 84%).

FIB used by a study was significantly associated with noncompliance at the source (p = 0.03) with mean noncompliance of TTC samples at 59% (95% CI: 41, 75%) and mean noncompliance of *E. coli* samples at 30% (95% CI: 16, 50%). There was a non-significant increase in the odds of source water noncompliance in longitudinal studies compared with cross sectional studies (OR = 3.9; 95% CI: 0.8, 19.4), but the estimate was based on only four longitudinal studies (Table 3).

Studies published after 2009 (median year of included studies) had significantly lower mean noncompliance rates compared to studies published in or before 2009 for both the source (p = 0.01) and HSW (p = 0.007) when all supply types are aggregated. Samples taken at the source were 4.0 (95% CI: 1.4, 11.6) times more likely to be noncompliant and samples taken from HSW were 3.9 times (95% CI: 1.5, 10.4) more likely to be noncompliant for studies published in or before 2009 when compared to studies published after 2009.

The heterogeneity of the studies was very high with a multivariate I_R^2 value of 0.91, indicating that 91% of total variance could be accounted for by between-study variance. There was no evidence of publication bias; the trim and fill test indicated no studies would need to be trimmed to create a symmetrical funnel plot (Figure 3).

Discussion

While the JMP estimated that 748 million people used unimproved water in 2012 (WHO/UNICEF 2014), several studies have modeled the global population drinking unsafe water through incorporation of water quality data at the source (Bain et al. 2014a; Onda et al. 2012). These refined estimates indicate that approximately 1.8 billion people lack access to safe drinking water (Onda et al. 2012), with 1.1 billion of these people using source water that is at least "moderate" risk (>10 *E. coli* or TTC per 100 mL) (Bain et al. 2014a). Onda et al. (2012) further corrected these estimates for sanitary inspection scores of water sources, concluding that a further 1.2 billion people use water from sources with multiple sanitary risks. Since we have found that HSW is substantially more likely to be contaminated than water at the source we suggest that even these refined estimates of the global population exposed to fecal contamination are likely to be overestimates.

Developed initially for its 2008 report, the JMP water ladder refined the concepts of "improved" and "unimproved" supplies (WHO/UNICEF 2008). It includes four rungs, descending from water piped on premises through other improved supplies, unimproved supplies (excluding surface water), and surface water. We found that source water from piped supplies was of significantly higher quality than that from other sources and this held true for HSW also, providing evidence to support the water ladder and promotion of piped water.

The point of collection for piped supplies – community standpipes, piped on plot and piped into the dwelling – is critical. The need for water storage is thought to be associated with distance to the collection point and reliability of the water source (Bain et al. 2014a). While the

JMP water ladder distinguishes between water piped on premises and other improved supplies, which include community standpipes, it was not possible to determine the location of the point of collection for many of the studies included in the review.

Piped water may be continuous, 24 hours per day seven days a week, predictably intermittent, or unreliable. Some studies in this review identify contamination related to non-continuous flow, including bacterial growth (Agard et al. 2002; Kumpel and Nelson 2013; Leiter et al. 2013). Others note the need to store piped water if supply is intermittent (Shaheed et al. 2014) which, as shown in our meta-analysis, may increase noncompliance. Finally, in our meta-analysis residual chlorine was only found for piped supplies (Table 2); however we were unable to analyze the effects of residual chlorine on water quality due to diverse reporting methods. To enable such an investigation, we recommend researchers report the presence of residual chlorine in drinking water samples tested for *E. coli*, especially those from piped supplies, using the WHO guideline values of 0.2 and 0.5 ppm (WHO 2011).

In our meta-analysis, we assessed the prevalence of non-compliant samples contaminated with *E. coli* or TTC, but not the public health impact of non-compliance. Our findings follow from Wright et al. (2004) as in general noncompliance is higher in HSW than at the source. However, we find that this relationship is modified by water supply type, with piped supplies having significantly lower odds of contamination at the source and in HSW than non-piped supplies. This finding may largely reflect the presence of residual chlorine in piped supplies that is uncommon amongst other supply types. Ensuring correct, consistent and continued usage of water treatment by households using non-piped supplies has proven to be challenging (Brown and Clasen 2012) and an area of active research (Ahuja et al 2010). While we anticipate that health impact will depend on different pathogens present in source and stored water, we argue that piped supplies may be safer, since they are likely to have fewer pathogens from both. We suggest that “leapfrogging” households up the water ladder, from unimproved sources to piped water could bring substantive health benefits. In addition, “leapfrogging” households could decrease abandoned investments, estimated at US$78 billion, which result from households passing stepwise through each rung of the water ladder, progressively abandoning their previous sources (Bain et al. 2013).

Leapfrogging households up the water ladder is unlikely to eliminate the need for water storage. Piped supplies which are unreliable or intermittent are common and necessitate water storage. It is widely believed that once households have predictable and reliable piped on premises, storage behaviors will decline. However Onda (2014) document continued storage behavior, the reasons for which are poorly understood, but may include anticipation of supply cuts (Onda 2014), taking advantage of the cooling activity of clay vessels (Klasen et al. 2012) or a refrigerator, convenience when a tap is on premises but not in the area for eating or drinking, and habit. To reduce storage practices, which may lead to higher noncompliance, a better understanding of the prevalence of and reasons for household water storage is necessary.

Contamination in HSW is widespread and it therefore appears credible that household water treatment and safe storage (HWTS) may have an interim role, especially in rural areas where access to piped water is less common in all regions. We were unable to assess the impact of HWTS on HSW quality due to lack of disaggregation in most studies. One study included samples untreated and reportedly boiled water stored in the household (Mertens et al. 1990). Reportedly-boiled HSW had 36-52% lower noncompliance depending on supply type than untreated HSW. Trevett et al. (2004), in which well water noncompliance was lower in HSW, suggested this apparent anomaly could have been due to inadvertent HSW treatment , such as dipping a ladle cleaned with bleach into the water. In general, however, both the efficacy of some HWTS methods and the determinants of consistent proper usage of HWTS remain inadequately understood and so the actual health benefits of HWTS remain unclear (Boisson et al. 2013, Brown and Clasen 2012; Enger et al. 2013).

One of the primary modes of contamination of HSW is contact with dirty hands and utensils (Psutka et al. 2011). Due to interaction with household hygiene, which lies in the private domain, stored water quality may not fall neatly into regulatory frameworks in the same manner as source water quality. The health sector should therefore play a key role in surveillance and policy-making to address the interaction between stored water quality and household hygiene (Rehfuess et al. 2009).

Methodological challenges: outcome level

While the dichotomous measure of compliance provides a snapshot of contamination, it contains very limited information about both the degree of contamination and its health significance. Presence/absence measurement was developed for monitoring where contamination was infrequent (Pipes et al. 1987), but because of its ease of use and lower cost, has been frequently applied where water contamination is common. When monitoring is infrequent and samples are often contaminated (on average we find 45% of samples at the source and 71% of samples from HSW were contaminated in this meta-analysis) the data become much less useful. While there is some evidence of a relationship between noncompliance and level of contamination (Bain et al. 2014a), a higher percentage of noncompliance does not necessarily indicate water is more highly contaminated, simply that contamination is widespread.

Methodological challenges: study level

One of the main methodological challenges of this review has been how to compare water quality at different sampling points from collection to consumption. We have chosen to analyze water quality compliance data at two sampling points, the source and HSW. Some studies (Klasen et al. 2012; Kumpel and Nelson 2013) suggest a more nuanced schema of the different points where contamination may occur, including the point of consumption, during transport, etcetera. Of particular importance is the point of consumption. HSW is not consumed directly, but is transferred to at least one other container or utensil before drinking. Thus, HSW quality data are likely to underestimate contamination of water as consumed and used in food preparation. Three studies in the qualitative synthesis sampled water from the drinking cup, a point closer to consumption. However, data were insufficient to generate pooled estimates of quality. While we present these data as two points on the pathway from collection to consumption, these group level data may not be connected at the level of individual samples.

In some studies, water points and households were sampled separately and data then aggregated by water supply and matched *post-hoc*. Within these, some studies sampled more HSW than sources where more than one household used a given source. In others, all sources in a study area were sampled, but HSW was only sampled for a fraction of sources. Aldana (2010) and Aliev et al. (2010), which include water supplies where noncompliance was lower in HSW than at the source fall into this latter category. While it is possible for quality to improve between source and storage (VanDerslice and Briscoe 1993), in both of these studies HSW was only sampled for 10% of sources and source and HSW quality were not linked, yet our analysis includes all HSW and source samples. Thus, we cannot draw inferences about individual samples.

Even for studies with a paired sampling design, with one water sample taken at a household's source and one from their water storage container, there are several reasons these samples cannot be considered "true" pairs. While households collect water and store it for a period of time, researchers often take samples closer together in time, often sampling HSW and then following up with source sampling after source identification by households.

Collecting data over a period of two years and visiting some households more than 10 times, Trevett et al. (2004) found high variation in contamination levels of samples taken over time from the same household, which they attributed to household behaviors. At the source, variability of contamination may be caused by factors such as seasonality (Kostyla et al. 2015). This potentially high variability in contamination at both the source and in HSW combined with the temporal dislocation in sampling noted above may introduce error. In addition, significant natural attenuation has been found in indicator organisms in stored water over time (see Levy et al. 2008 for example), while longer storage time also presents more opportunities for contamination. Variation in how long water has been stored in the household before sampling is thus likely to be a source of heterogeneity in water quality data both within and between studies along with other factors such as temperature.

Even if sampling and timing could be rigorously controlled and examined, the notion of true paired samples is problematic because it assumes a one-to-one ratio between source and storage. However, classification of a 'primary' source is a simplification because households often use multiple sources (Shaheed et al. 2014). There is also a potential for misreporting of sources by households, where household members claim to use improved sources rather than unimproved with consequences for apparent level of contamination.

Heterogeneity of studies was very high. In addition to the methodological issues for measuring and comparing noncompliance listed above, this high level of variation reflects the fact that water sources and household water storages are located within complex systems. Some aspects of these complex systems which impact noncompliance are touched upon in the included studies For example, while studies may report the proportion of participants treating water, FIB data disaggregated by household treatment were rarely reported, preventing meta-analysis of household treatment. Rural/urban geography is reported in the studies; however population densities for what is considered rural and urban may vary widely between and even within countries (Christenson et al. 2014). Yet our knowledge of these complex systems is limited and this high variability, which remains after exploring known confounders and effect modifiers, is evidence that we have a lack of understanding the system in which contamination at the source and in HSW occurs.

Methodological challenges: review level

In their original review on source versus stored water quality, Wright et al. (2004) collapse noncompliance data on quality at the source and quality in HSW into a single odds ratio. The associated loss of information and dimensionality inhibits analysis of which covariates are relevant at each sampling point. Bivariate techniques were developed for meta-analyses of sensitivity and specificity of medical diagnostic tests in order to avoid the loss of data and dimensionality associated with the use of the diagnostic odds ratio (Menke 2010; Reitsma et al. 2005). Assuming a bivariate normal distribution for noncompliance at the source and in HSW, we apply the same analytical technique to water quality.

While bivariate techniques are becoming more common in meta-analyses of medical diagnostic tests, they have not been frequently used outside that field. Using both bivariate and univariate methods has allowed us to avoid data and dimensionality loss of converting noncompliance at the source and in HSW to a simple odds ratio while also exploring the relative contamination at both sampling points. Testing for heterogeneity in bivariate and multivariate meta-analysis is a new but rapidly developing field, while bivariate methods for exploring publication bias have not yet been developed, necessitating use of univariate tests. The most recent literature suggests that the most appropriate application of a univariate method to test for publication bias is to apply the trim and fill method to the log odds ratio of the two variables (Bürkner and Doebler 2014).

One assumption of this meta-analysis is that in studies that have data on multiple water supply types the contamination of each of these supplies is independent of the others. Since water quality is affected by environmental sanitation and other community level factors, this assumption may not always hold true. In addition, this method gives more weight to studies with a higher number of water supply types studied. However, since the majority of the studies explore only one supply type and only seven of 45 studies include three or more supply types it is not feasible to group water supplies by study for analysis.

Finally, this review was limited by the number of studies identified for some supply types and the quality of reporting. In particular, infrequent and inconsistent reporting of residual chlorine meant that we were not able to determine whether this was one of the main reasons piped supplies were less likely to be contaminated. Lack of a definition of HSW may have contributed to high heterogeneity within the stored water quality results. In addition, it is possible that we have missed studies due to lack of standard terminology. We suggest that adopting a definition of HSW will enable better comparisons across studies and contexts.

Conclusions

We find substantive evidence for deterioration in water quality between source and stored water. As such, estimates of the global population drinking safe water, even those that account for water quality and sanitary inspection, are likely to be overestimates due to contamination during collection, transport or storage. We propose that monitoring of drinking water quality should occur at both the source and in HSW. We find that piped water is significantly less likely to be contaminated than other water supply types, both at the source and in HSW and suggest that a shift toward piped supplies will lead to both improved quality and safety of drinking water in the household. While previous development policies have focused on extending a basic level of service to all, we suggest that future development policies, such as the Sustainable Development Goals need to incorporate goals of moving people up the water ladder. While HWTS may have a role to play in the short term, improving source water quality, particularly of piped sources is likely to lead to improved quality at both the source and in HSW. In particular, a consistent supply of high quality piped water on premises is likely to lead to the highest quality drinking water, even if storage continues. We see a role for the health sector in surveillance and policy making to address the interaction between stored water quality and household hygiene. In order to evaluate the success of future development policies in providing safe drinking water, future studies should seek to move beyond presence/absence measures to report FIB or even pathogen counts and variances, in addition to recording and reporting residual chlorine using the WHO guideline values of 0.2 and 0.5 ppm (WHO 2011).

References

Abdellah AM, Abdel-magid HM, Yahia NA. 2012. Assessment of Drinking Water Microbial Contamination in Al-Butana Region of Sudan. J. Appl. Sci. 12:856–862; doi:DOl: 10.3923/jas.2012.856.862.

Agard L, Alexander C, Green S, Jackson M, Patel S, Adesiyun A. 2002. Microbial quality of water supply to an urban community in Trinidad. J Food Prot. 65(8):1297–1303.

Ahuja, A., Kremer, M., & Zwane, A. P. 2010. Providing safe water: Evidence from randomized evaluations. Annu Rev Resour Economics, 2(1), 237–256.

Aldana J. 2010. Rapid assessment of drinking-water quality in the Republic of Nicaragua: country report of the pilot project implementation in 2004-2005. Available: http://www.wssinfo.org/fileadmin/user_upload/resources/RADWQ_Nicaragua.pdf [accessed 29 March 2015].

Aliev S, Shodmonov P, N B, Schmoll O. 2010. Rapid assessment of drinking-water quality in the Republic of Tajikistan: country report of the pilot project implementation in 2004–2005. Available: http://www.wssinfo.org/fileadmin/user_upload/resources/RADWQ_Tajikistan.pdf [accessed 29 March 2015].

Austin CJ. Chlorinating household water in the Gambia. In: Proceedings of the 20th WEDC Conference Colombo, Sri Lanka, 1993. Available: http://wedc.lboro.ac.uk/resources/conference/19/austin.pdf [accessed 29 March 2015].

Bain R, Luyendijk R, Bartram J. 2013. Universal access to drinking water: the role of aid.WIDER Working Paper. Available: http://www.econstor.eu/bitstream/10419/93683/1/768673054.pdf [accessed 26 April 2015].

Bain R, Cronk R, Hossain R, Bonjour S, Onda K, Wright J, et al. 2014a. Global assessment of exposure to faecal contamination through drinking water based on a systematic review. Trop Med Int Health 19(8):917–927. doi: 10.1111/tmi.12334.

Bain R, Cronk R, Wright J, Yang H, Slaymaker T, Bartram J. 2014b. Fecal contamination of drinking-water in low-and middle-income countries: a systematic review and meta-analysis. PLoS Med 11(5):e1001644.

Baker KK, Sow SO, Kotloff KL, Nataro JP, Farag TH, Tamboura B, et al. 2013. Quality of Piped and Stored Water in Households with Children Under Five Years of Age Enrolled in the Mali Site of the Global Enteric Multi-Center Study (GEMS). Am J Trop Med Hyg 89(2):214.

Bartram J, Fisher MB, Brocklehurst C, Luyendijk R, Hossain R, Wardlaw T, et al. 2014. Global Monitoring of Water Supply and Sanitation: History, Methods and Future Challenges. Int J Environ Res Public Health 11(8):8137–8165.

Boisson S, Stevenson M, Shapiro L, Kumar V, Singh LP, Ward D, et al. 2013. Effect of household-based drinking water chlorination on diarrhoea among children under five in Orissa, India: a double-blind randomised placebo-controlled trial. PLoS Med 10(8):e1001497.

Bradley DJ, Bartram JK. 2013. Domestic water and sanitation as water security: monitoring, concepts and strategy. Philosophical transactions Series A, Mathematical, physical, and engineering sciences 371(2002):20120420.

Brown J, Clasen T. 2012. High adherence is necessary to realize health gains from water quality interventions. Plos One 7(5):e36735.

Bürkner PC, Doebler P. 2014. Testing for publication bias in diagnostic meta-analysis: a simulation study. Stat Med. 33(18):3061–3077

Chemuliti JK, Gathura PB, Kyule MM, Njeruh FM. 2002. Bacteriological qualities of indoor and out-door drinking water in Kibera sub-location of Nairobi, Kenya. East Afr Med J 79(5):271–273.

Christenson E, Bain R, Wright J, Aondoakaa S, Hossain R, Bartram J. 2014. Examining the influence of urban definition when assessing relative safety of drinking-water in Nigeria. Sci Total Environ 490:301–312.

Chung C. 2011. The challenges of a water system management handover in Eastern Ethiopia - from the United Nations Refugee Agency to a local community. Cambridge: Massachusetts Institute of Technology.

Clasen TF, Bastable A. 2003. Faecal contamination of drinking water during collection and household storage: the need to extend protection to the point of use. J Water Health 1(3):109–115.

Clasen TF, Cairncross S. 2004. Household water management: refining the dominant paradigm. Trop Med Int Health 9(2):187–191.

Cronin AA, Breslin N, Gibson J, Pedley S. 2006. Monitoring source and domestic water quality in parallel with sanitary risk identification in northern Mozambique to prioritise protection interventions. J Water Health 4:333–345.

de Sá LLC, Jesus IMd, Santos ECO, Vale ER, Loureiro ECB, Sá EVd. 2005. Microbiological quality of drinking water in two areas following sanitation interventions-Belém, Pará State, Brazil. Epidemiologia e Serviços de Saúde 14(3):171–180.

Elala D, Labhasetwar P, Tyrrel SF. 2011. Deterioration in water quality from supply chain to household and appropriate storage in the context of intermittent water supplies. Wa Sci Technol 11(4):400–408.

Enger KS, Nelson KL, Clasen T, Rose JB, Eisenberg JNS. 2012. Linking quantitative microbial risk assessment and epidemiological data: informing safe drinking water trials in developing countries. Environ Sci Technol 46(9):5160–7; doi.org/10.1021/es204381e.

Eshcol J, Mahapatra P, Keshapagu S. 2009. Is fecal contamination of drinking water after collection associated with household water handling and hygiene practices? A study of urban slum households in Hyderabad, India. J Water Health 7(1):145–154.

Fiore MM, Minnings K, Fiore LD. 2010. Assessment of biosand filter performance in rural communities in southern coastal Nicaragua: an evaluation of 199 households. Rural and Remote Health 10(3):1483.

Firth J, Balraj V, Muliyil J, Roy S, Rani LM, Chandresekhar R, et al. 2010. Point-of-use interventions to decrease contamination of drinking water: a randomized, controlled pilot study on efficacy, effectiveness, and acceptability of closed containers, Moringa oleifera, and in-home chlorination in rural South India. Am. J. Trop. Med. Hyg. 82:759–65; doi:10.4269/ajtmh.2010.09-0206.

Genthe B, Strauss N, Seager J, Vundule C, Maforah F, Kfir R. 1997. The Effect of Type of Water Supply on Water Quality in a Developing Community in South Africa. Water Sci. Technol. 35:35–40.

Gerba CP, Rose JB, Haas CN. 1996. Sensitive populations: who is at the greatest risk? Int J Food Microbiol 30(1-2):113–123.

Gundry S, Wright J, Conroy R. 2004. A systematic review of the health outcomes related to household water quality in developing countries. J Water Health 2(1):1–13.

Handzel T. 1998. The effect of improved drinking water quality on the ridk of diarrheal disease in an urban slum of Dhaka, Bangladesh: A home chlorination intervention trial. University of North Carolina at Chapel Hill.

Holm RH. 2012. Survey of water quality in Northern Malawi, Africa, and communication of human health risk. Washington State University.

Hoque BAHallman K, Levy J, Bouis H, Ali N, Khan F, et al. 2006. Rural drinking water at supply and household levels: quality and management. Int J Hyg Environ Health 209:451–60; doi:10.1016/j.ijheh.2006.04.008.

Jackson D, White IR, Riley RD. 2012. Quantifying the impact of between-study heterogeneity in multivariate meta-analyses. Stat Med 31(29):3805–3820; doi: 10.1002/sim.5453.

Jagals P, Barnard TG, Mokoena MM, Ashbolt N, Roser DJ. 2013. Pathogenic Escherichia coli in rural household container waters. Water Sci Technol 67(6):1230–1237.

Jagals P, Bokako TC, Grabow WOK. 1999. Changing Consumer Water-Use Patterns and Their Effect on Microbiological Water Quality as a Result of an Engineering Intervention. Water S A 25(3):297–297.

Jagals P, Grabow WOK, Williams E. 1997. The effects of supplied water quality on human health in an urban development with limited basic subsistence facilities. Water Sa 23(4):373–378.

Kanyerere T, Levy J, Xu Y, Saka J. 2012. Assessment of microbial contamination of groundwater in upper Limphasa River catchment, located in a rural area of northern Malawi. Water SA (Pretoria) 38(4):581–595.

Khush R, London A, Colford J, Arnold B, Balakrishnan K, Ramaswamy P. 2009. Evaluating the Sustainability and Impacts of Water, Sanitation & Hygiene Interventions. Aquaya, San Francisco, CA.

Klasen S, Lechtenfeld T, Meier K, Rieckmann J. 2012. Benefits trickling away: the health impact of extending access to piped water and sanitation in urban Yemen. Journal of Development Effectiveness 4(4):537–565.

Kostyla C, Bain R, Cronk R, Bartram J. 2015. Seasonal variation of fecal contamination in drinking water sources in developing countries: A systematic review. Sci Total Environ 514:333–343.

Kremer M, Leino J, Miguel E, Zwane a. P. 2011. Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions. Q. J. Econ. 126:145–205; doi:10.1093/qje/qjq010.

Kumpel E, Nelson KL. 2013. Comparing Microbial Water Quality in an Intermittent and Continuous Piped Water Supply. Water Res 47(14):5176–5188.

Lacey S, Lopez R, Frangos C, Khodadoust A. 2011. Water Quality Degradation after Water Storage at Household Level in a Piped Water System in Rural Guatemala. International Journal for Service Learning in Engineering, Humanitarian Engineering and Social Entrepreneurship 6(1):118–129.

Levy K, Nelson KL, Hubbard A, Eisenberg JNS. 2008. Following the Water: A Controlled Study of Drinking Water Storage in Northern Coastal Ecuador. Environ Health Perspect 116(11):1533–1540.

Magrath J. 2006. Towards sustainable water-supply solutions in Rural Sierra Leone: a pragmatic approach, using comparisons with Mozambique. Oxford:Oxfam/WaterAid. Available: http://policy-practice.oxfam.org.uk/publications/towards-sustainable-water-supply-solutions-in-rural-sierra-leone-a-pragmatic-ap-112384 [accessed 19 April 2015].

Mazengia E, Chidavaenzi MT, Bradley M, Jere M, Nhandara C, Chigunduru D, et al. 2002. Effective and culturally acceptable water storage in Zimbabwe: Maintaining the quality of water abstracted from upgraded family wells. J Environ Health 64(8):15–18.

Menke J. 2010. Bivariate random-effects meta-analysis of sensitivity and specificity with SAS PROC GLIMMIX. Methods Inf Med 49(1):54–62.

Mertens TE, Fernando MA, Marshall TFDC, Kirkwood BR, Cairncross S, Radalowicz A. 1990. Determinants of water quality availability and use in Kurunegala Sri Lanka. Trop Med Parasitol 41(1):89–97.

Mintz E, Bartram J, Lochery P, Wegelin M. 2001. Not just a drop in the bucket: expanding access to point-of-use water treatment systems. Am J Public Health 91(10):1565–1570.

Oloruntoba EO, Sridhar MK. 2007. Bacteriological quality of drinking water from source to household in Ibadan, Nigeria. Afr J Med Med Sci 36(2):169–175.

Onda, K. 2014. Intermittent vs. continuous supply: What benefits do households actually receive? Evidence from two cities in India. [MSPH Technical Report]. Chapel Hill, NC:University of North Carolina, Chapel Hill. Available: https://cdr.lib.unc.edu/record/uuid:f6e8d52f-bd32-4191-9f0b-cdee3f1272de [accessed 29 March 2015].

Onda K, LoBuglio J, Bartram J. 2012. Global access to safe water: accounting for water quality and the resulting impact on MDG progress. Int J Environ Res Public Health 9(3):880–894.

Pakula, L. 2008. Correlation, Regression, Bivariate Normal. Available: http://www.math.uri.edu/~pakula/452webs8/regress.pdf [accessed 7 July 2014].

Pickering AJ, Davis J, Walters SP, Horak HM, Keymer DP, Mushi D, et al. 2010. Hands, Water, and Health: Fecal Contamination in Tanzanian Communities with Improved, Non-Networked Water Supplies. Environ Sci Technol 44(9):3267–3272.

Pipes WO, Mueller K, Troy M, Minnigh HA. 1987. Frequency-of-Occurrence Monitoring for Coliform Bacteria in Small Water Systems. Journal (American Water Works Association) 79(11):59–63.

Platenburg RJPM, Zaki M. 1993. Patterns of Water Quality in Rural Areas of Assyut Governorate, Egypt. Water Science & Technology 27(9):55–55.

Potgieter N, Becker PJ, Ehlers MM. 2009. Evaluation of the CDC safe water-storage intervention to improve the microbiological quality of point-of-use drinking water in rural communities in South Africa. Water Sa 35(4):505–516.

Psutka R, Peletz R, Michelo S, Kelly P, Clasen T. 2011. Assessing the microbiological performance and potential cost of boiling drinking water in urban Zambia. Environ Sci Technol 45(14):6095–6101.

Quick RE, Kimura A, Thevos A, Tembo M, Shamputa I, Hutwagner L, et al. 2002. Diarrhea prevention through household-level water disinfection and safe storage in Zambia. American J Trop Med Hyg 66(5):584–589.

Rehfuess EA, Bruce N, Bartram JK. 2009. More health for your buck: health sector functions to secure environmental health. Bull World Health Organ 87(11):880–882; doi.org/10.2471/BLT.08.059865

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. 2005. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 58(10):982–990.

Rosa G, Miller L, Clasen T. 2010. Microbiological Effectiveness of Disinfecting Water by Boiling in Rural Guatemala. Am J Trop Med Hyg 82(3):473–477.

Shaheed A, Orgill J, Ratana C, Montgomery MA, Jeuland MA, Brown JM. 2014. Water quality risks of "improved" water sources: evidence from Cambodia. Trop Med Int Health, 19(2):186–194.

Shar AH, Kazi YF, Kanhar NA, Soomro IH, Zia SM, Ghumro PB. 2010. Drinking water quality in Rohri City, Sindh, Pakistan. African Journal of Biotechnology 9(42):7102–7107.

Shrestha S, S. Malla S, Aihara Y, Kondo N, Nishida K. 2013. Water Quality at Supply Source and Point of Use in the Kathmandu Valley. J. Water Environ. Technol. 11:331–340; doi:10.2965/jwet.2013.331.

Simango C, Dindiwe J, Rukure G. 1992. Bacterial contamination of food and household stored drinking water in a farmworker community in Zimbabwe. Cent Afr J Med 38(4):143–149.

Stoler J, Fink G, Weeks JR, Otoo RA, Ampofo JA, Hill AG. 2012. When urban taps run dry: Sachet water consumption and health effects in low income neighborhoods of Accra, Ghana. Health Place 18(2):250–262.

Sutton S, Butterworth J, Mekonta L. 2012. A hidden resource: household-led rural water supply in Ethiopia. IRC International Water and Sanitation Centre, the Netherlands. Available: http://www.ircwash.org/resources/hidden-resource-household-led-rural-water-supply-ethiopia

Tabor M, Kibret M, Abera B. 2011. Bacteriological and physicochemical quality of drinking water and hygiene-sanitation practices of the consumers in bahir dar city, ethiopia. Ethiop J Health Sci 21(1):19–26.

Tadesse D, Desta A, Geyid A, Girma W, Fisseha S, Schmoll O. 2010. Rapid assessment of drinking-water quality in the Federal Democratic Republic of Ethiopia: country report of the pilot project implementation in 2004–2005. Available: http://www.wssinfo.org/fileadmin/user_upload/resources/RADWQ_Ethiopia.pdf [accessed 19 April 2015].

Trevett AF, Carter RC, Tyrrel SF. 2004. Water quality deterioration: A study of household drinking water quality in rural Honduras. Int J Environ Health Res 14(4):273–283.

Trevett AF, Carter RC, Tyrrel SF. 2005. The importance of domestic water quality management in the context of faecal – oral disease transmission. J Water Health, 3(3):259–270.

VanDerslice J, Briscoe J. 1993. All coliforms are not created equal: A comparison of the effects of water source and in-house water contamination on infantile diarrheal disease. Water Resour Res 29(7):1983–1995.

WHO (World Health Organization). 2011. Guidelines for drinking-water quality. 4th ed. Geneva: World Health Organization. Available: http://whqlibdoc.who.int/publications/2011/9789241548151_eng.pdf [accessed 22 April 2015].

WHO/UNICEF (World Health Organization / United Nations Children's Fund). 2008. Progress on Drinking-Water and Sanitation: Special Focus on Sanitation. Geneva, Switzerland:WHO/UNICEF. Available: http://www.wssinfo.org/fileadmin/user_upload/resources/1251794333-JMP_08_en.pdf [accessed 26 April 2015].

WHO/UNICEF (World Health Organization / United Nations Children's Fund). 2012. Progress on Sanitation and Drinking-Water: 2012 Update. Geneva, Switzerland:WHO/UNICEF. Available: http://www.wssinfo.org/fileadmin/user_upload/resources/JMP-report-2012-en.pdf [accessed 26 April 2015].

WHO/UNICEF (World Health Organization / United Nations Children's Fund). 2014. Progress on Drinking-Water and Sanitation: 2014 Update. Geneva, Switzerland:WHO/UNICEF. Available: http://www.wssinfo.org/fileadmin/user_upload/resources/JMP_report_2014_webEng.pdf [accessed 26 April 2015].

World Bank. 2013. Country and Lending Groups. World Bank Group, Washington, DC. Available: http://data.worldbank.org/data-catalog/world-development-indicators [accessed 1 December 2013].

Wright J, Gundry S, Conroy R. 2004. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Trop Med Int Health 9(1):106–117.

Zhou Y, Dendukuri N. 2014. Statistics for quantifying heterogeneity in univariate and bivariate meta-analyses of binary data: The case of meta-analyses of diagnostic accuracy. Stat Med 33(16):2701–2717.

Table 1: Included studies.

Study	MDG region	Country income level	Setting	Indicator	Study design	Study quality	Random selection	Publication year	Measure of central tendency
Abdellah et al. 2012	sub-Saharan Africa	Low	Rural	EC	Cross sectional	Low	No	After 2009	Yes
Agard et al. 2002	Latin America and the Caribbean	High	Urban	EC	Cross sectional	Medium	Yes	In or before 2009	No
Aldana 2010	Latin America and the Caribbean	Lower-middle	Both	TTC	Cross sectional	High	Yes	After 2009	No
Aliev et al. 2010	Caucuses and Central Asia	Low	Both	TTC	Cross sectional	High	Yes	After 2009	Yes
Austin 1993	sub-Saharan Africa	Low	Rural	TTC	Intervention	Low	No	In or before 2009	Yes
Baker et al. 2013	sub-Saharan Africa	Low	Urban	EC	Cross sectional	High	Yes	After 2009	Yes
Chemuliti et al. 2002	sub-Saharan Africa	Low	Urban	TTC	Cross sectional	Medium	No	In or before 2009	Yes
Chung 2011	sub-Saharan Africa	Low	Both	EC	Cross sectional	Low	No	After 2009	Yes
Cronin et al. 2006	sub-Saharan Africa	Low	Urban	TTC	Longitudinal	Low	No	In or before 2009	Yes
de Sá et al. 2005	Latin America and the Caribbean	Upper-middle	Urban	TTC	Longitudinal	High	Yes	In or before 2009	Yes
Elala et al. 2011	South Asia	Lower-middle	Urban	TTC	Cross sectional	Medium	No	After 2009	No
Eshcol et al. 2009	South Asia	Lower-middle	Urban	EC	Cross sectional	Medium	No	In or before 2009	Yes
Fiore et al. 2010[1]	Latin America and the Caribbean	Lower-middle	Rural	EC	Intervention	Low	No	After 2009	No
Firth et al. 2010	South Asia	Lower-middle	Rural	TTC	Intervention	Low	No	After 2009	Yes
Genthe et al. 1997	sub-Saharan Africa	Upper-middle	Urban	EC	Case-control	Low	No	In or before 2009	Yes
Handzel 1998	South Asia	Low	Urban	EC	Intervention	High	Yes	In or before 2009	Yes
Holm 2012[2]	sub-Saharan Africa	Low	Urban	EC	Cross sectional	Medium	No	After 2009	Yes

Study	**MDG region**	**Country income level**	**Setting**	**Indicator**	**Study design**	**Study quality**	**Random selection**	**Publication year**	**Measure of central tendency**
Hoque et al. 2006	South Asia	Low	Rural	TTC	Cross sectional	High	No	In or before 2009	Yes
Jagals et al 2013	sub-Saharan Africa	Upper-middle	Rural	EC	Cross sectional	Low	No	After 2009	No
Jagals et al. 1999	sub-Saharan Africa	Upper-middle	Urban	TTC	Intervention	Medium	No	In or before 2009	Yes
Jagals et al. 1997	sub-Saharan Africa	Upper-middle	Urban	TTC	Cross sectional	Medium	Yes	In or before 2009	Yes
Kanyerere et al. 2012	sub-Saharan Africa	Low	Rural	EC	Cross sectional	Medium	Yes	After 2009	Yes
Khush et al. 2009	South Asia	Lower-middle	Rural	EC	Intervention	Low	No	In or before 2009	Yes
Klasen et al. 2012[2]	Western Asia	Lower-middle	Urban	EC	Intervention	Low	Yes	After 2009	Yes
Kremer et al. 2011	sub-Saharan Africa	Low	Rural	EC	Intervention	Medium	Yes	After 2009	Yes
Kumpel and Nelson 2013[2]	South Asia	Lower-middle	Urban	EC	Longitudinal	Medium	No	After 2009	Yes
Lacey et al. 2011	Latin America and the Caribbean	Lower-middle	Rural	EC	Cross sectional	High	Yes	After 2009	Yes
Magrath 2006	sub-Saharan Africa	Low	Rural	TTC	Cross sectional	Low	No	In or before 2009	Yes
Mazengia et al. 2002	sub-Saharan Africa	Low	Rural	TTC	Intervention	Medium	No	In or before 2009	Yes
Mertens et al. 1990	South Asia	Lower-middle	Rural	TTC	Case-control	High	No	In or before 2009	Yes
Oloruntoba and Sridhar 2007	sub-Saharan Africa	Lower-middle	Urban	EC	Cross sectional	High	Yes	In or before 2009	Yes
Pickering et al. 2010	sub-Saharan Africa	Low	Urban	EC	Cross sectional	High	Yes	After 2009	Yes
Platenburg and Zaki 1993	North Africa	Lower-middle	Rural	TTC	Intervention	Low	No	In or before 2009	Yes
Potgieter et al. 2009[2]	sub-Saharan Africa	Upper-middle	Rural	EC	Intervention	Medium	Yes	In or before 2009	Yes
Quick et al. 2002	sub-Saharan Africa	Lower-middle	Urban	EC	Intervention	High	Yes	In or before	No

Study	**MDG region**	**Country income level**	**Setting**	**Indicator**	**Study design**	**Study quality**	**Random selection**	**Publication year**	**Measure of central tendency**
								2009	
Rosa et al. 2010[1]	Latin America and the Caribbean	Lower-middle	Rural	TTC	Cross sectional	Medium	No	After 2009	Yes
Shaheed et al. 2014	South East Asia	Low	Urban	EC	Cross sectional	Medium	Yes	After 2009	Yes
Shar et al. 2010	South Asia	Lower-middle	Urban	EC	Cross sectional	Low	No	After 2009	Yes
Shrestha et al. 2013	South Asia	Low	Rural	EC	Cross sectional	Low	No	After 2009	No
Simango et al. 1992	sub-Saharan Africa	Low	Rural	EC	Cross sectional	Medium	No	In or before 2009	No
Stoler et al. 2012	sub-Saharan Africa	Lower-middle	Urban	EC	Cross sectional	Low	No	After 2009	Yes
Sutton et al. 2012	sub-Saharan Africa	Low	Rural	TTC	Cross sectional	Low	No	After 2009	Yes
Tabor et al. 2011	sub-Saharan Africa	Low	Urban	TTC	Cross sectional	Medium	Yes	After 2009	No
Tadesse et al. 2010	sub-Saharan Africa	Low	Both	TTC	Cross sectional	High	No	After 2009	Yes
Trevett et al. 2004	Latin America and the Caribbean	Lower-middle	Rural	TTC	Longitudinal	Medium	No	In or before 2009	Yes

[1]Study excluded from meta-regression because all samples were reportedly treated in the household.

[2]Study excluded from meta-regression because water was sampled from the drinking cup rather from the household water storage container.

Table 2: Characteristics of included studies.

Variable	Number of studies n (%)	Water supply types evaluated n (%)	Source water samples n (%)	Household stored water samples n (%)
Total	45	65	10934	12523
MDG region				
Caucuses and Central Asia	1 (2)	2 (3)	1319 (12)	119 (1)
Latin America and the Caribbean	7 (16)	10 (15)	2247 (21)	1301 (10)
North Africa	1 (2)	1 (2)	189 (2)	183 (1)
South Asia	10 (22)	13 (20)	2703 (25)	5904 (47)
South East Asia	1 (2)	1 (2)	124 (1)	124 (1)
sub-Saharan Africa	24 (53)	36 (55)	3854 (35)	4391 (35)
Western Asia	1 (2)	2 (3)	498 (5)	501 (4)
Country income level				
High	1 (2)	1 (2)	81 (1)	104 (1)
Upper-middle	6 (13)	7 (11)	504 (5)	866 (7)
Lower-middle	17 (38)	25 (38)	5288 (48)	7468 (60)
Low	21 (47)	32 (49)	5061 (46)	4085 (33)
Setting				
Rural	20 (44)	27 (42)	3466 (32)	8181 (65)
Urban	21 (47)	28 (43)	3035 (28)	3871 (31)
Both	4 (9)	10 (15)	4433 (41)	471 (4)
FIB[1]				
E. coli	25 (56)	31 (48)	3659 (33)	8177 (65)
TTC[2]	20 (44)	34 (52)	7275 (67)	4346 (35)
Study design				
Case-control	2 (4)	4 (6)	1068 (10)	1687 (13)
Cross sectional	27 (60)	40 (62)	6482 (59)	3293 (26)
Intervention	12 (27)	14 (22)	2286 (21)	6123 (49)
Longitudinal	4 (9)	7 (11)	1098 (10)	1420 (11)
Study quality				
Low	16 (36)	23 (35)	2440 (22)	4831 (39)
Medium	17 (38)	18 (28)	1931 (18)	3821 (31)
High	12 (27)	24 (37)	6563 (60)	3871 (31)
Random selection[3]				
No	28 (62)	37 (57)	4880 (45)	8109 (65)
Yes	17 (38)	28 (43)	6054 (55)	4414 (35)
Publication year				
In or before 2009	21 (47)	30 (46)	3887 (36)	7278 (58)
After 2009	24 (53)	35 (54)	7047 (64)	5245 (42)

Variable	Number of studies n (%)	Water supply types evaluated n (%)	Source water samples n (%)	Household stored water samples n (%)
Measure of central tendency[4]				
No	9 (20)	12 (18)	1985 (18)	790 (6)
Yes	36 (80)	53 (82)	8949 (82)	11733 (94)
Water supply type				
Improved				
Piped		31 (48)	5425 (50)	6316 (50)
Borehole		12 (18)	1749 (16)	1736 (14)
Protected well		8 (12)	1867 (17)	1724 (14)
Protected Spring		2 (3)	654 (6)	59 (0)
Unimproved				
Unprotected well		4 (6)	316 (3)	220 (2)
Unprotected spring		1 (2)	193 (2)	1445 (12)
Tanker truck		1 (2)	211 (2)	212 (2)
Unspecified				
Unspecified well		6 (9)	519 (5)	811 (6)
Treatment of piped sources[5]				
Reported chlorination, residual chlorine assessed		8 (26)	1416 (26)	937 (15)
Reported chlorination, residual chlorine NOT assessed		6 (19)	578 (11)	943 (15)
Inconsistent chlorination, residual chlorine not assessed		6 (19)	2026 (37)	407 (6)
Not treated		2 (6)	240 (4)	216 (3)
Not reported		9 (29)	1165 (21)	3813 (60)

[1]FIB stands for Fecal Indicator Bacteria

[2]TTC stands for thermotolerant coliforms

[3]Random selection was assessed through the question "Was sampling randomized over a given study area or population?"

[4]Studies were considered to include a measure of central tendency if they reported mean, geometric mean and/or median level of contamination by water supply type.

[5]The total n for this set of characteristics is 31, corresponding to the number of studies that reported information on piped supplies. No other supply types were reported to be chlorinated or to have been tested for residual chlorine.

Table 3: Mean proportion of noncompliant samples from between studies meta-regression.

Covariate	**Source: n water samples**	**Source: mean noncompliance (95% CI)**	**Source: p-value**	**HSW: n water samples**	**HSW: mean noncompliance (95% CI)**	**HSW: p-value**
Unadjusted	9198	46 (33, 60)	0.59	10557	75 (64, 84)	<.001
Water supply type			<.001			0.03
Improved						
Piped	4195	25 (15, 40)		4801	62 (44, 77)	
Borehole	1749	43 (22, 66)		1736	83 (63, 93)	
Protected well	1753	70 (41, 89)		1666	80 (51, 94)	
Protected Spring	654	35 (5, 84)		59	46 (6, 91)	
Unimproved						
Unprotected well	286	94 (69, 99)		190	94 (65, 99)	
Unprotected spring	654	5 (0, 61)		1445	14 (0, 83)	
Unspecified						
Unspecified well	368	91 (68, 98)		660	94 (77, 99)	
MDG region			0.97			0.71
Caucuses and Central Asia	1319	14 (1, 78)	0.91	119	10 (1, 61)	0.17
Latin America and the Caribbean	1872	54 (20, 84)		944	67 (36, 88)	
North Africa	189	61 (2, 99)		183	89 (17, 100)	
South Asia	2166	51 (22, 79)		5308	85 (65, 94)	
South East Asia	124	53 (1, 99)		124	82 (10, 99)	
sub-Saharan Africa	3528	44 (26, 63)		5308	76 (61, 86)	
Country income level			0.03			0.23
High	81	33 (1, 96)		104	67 (5, 99)	
Upper-middle	322	9 (1, 43)		442	39 (9, 81)	
Lower-middle	3878	68 (47, 84)		6014	84 (68, 92)	
Low	4917	39 (23, 56)		3997	73 (57, 84)	
Setting			0.44			0.67
Rural	2961	55 (32, 77)		7756	79 (62, 89)	
Urban	1804	42 (21, 66)		2330	83 (67, 92)	
FIB			0.03			0.42
E. coli	2427	30 (16, 50)		6889	71 (52, 84)	
TTC	6771	59 (41, 75)		3668	79 (65, 88)	
Study design			0.17			0.21
Case-control	1068	52 (12, 89)		1687	70 (26, 94)	
Cross sectional	6199	72 (37, 92)		2957	91 (72, 98)	
Intervention	1370	69 (30, 92)		5089	84 (52, 96)	
Longitudinal	591	36 (22, 52)		824	69 (54, 81)	
Study quality			0.44			0.66
Low	1791	44 (24, 67)		4179	70 (49, 85)	
Medium	844	31 (12, 62)		2507	72 (45, 89)	

Covariate	Source: n water samples	Source: mean noncompliance (95% CI)	Source: p-value	HSW: n water samples	HSW: mean noncompliance (95% CI)	HSW: p-value
High	6563	55 (34, 74)		3871	80 (64, 90)	
Random selection			0.61			0.99
No	3687	49 (31, 68)		6754	75 (60, 86)	
Yes	5511	42 (24, 63)		3803	75 (58, 87)	
Publication year			0.01			0.007
In or before 2009	3591	64 (45, 79)		6796	86 (75, 93)	
After 2009	5607	31 (18, 48)		3761	61 (45, 76)	
Measure of central tendency			0.54			0.30
No	1834	38 (15, 68)		639	64 (36, 85)	
Yes	7364	48 (33, 64)		9918	78 (66, 86)	

Data were derived from bivariate random effects regression of the number of noncompliant samples out of the total number of samples. The unadjusted model contained fixed effects of source noncompliance and HSW noncompliance. All adjusted models included one covariate of interest an interaction with source noncompliance and an interaction term with HSW noncompliance. Model estimates were back transformed to derive estimates of mean compliance on the scale of the original data. P-values were calculated using the type III test of fixed effects.

Table 4: Odds ratios for microbial non-compliance, comparing source, study setting and study design characteristics, calculated for source and household stored water samples.

Contrast	Source: OR (95% CI)	Source: p-value	HSW: OR (95% CI)	HSW: p-value
Unimproved vs improved supplies	6.5 (1.0, 43.9)	0.06	2.0 (0.3, 14.)	0.48
Improved supplies: piped vs other	0.3 (0.1, 0.8)	0.01	0.4 (0.2, 1.2)	0.12
All supplies: piped vs other	0.2 (0.1, 0.5)	0.002	0.3 (0.2, 0.8)	0.03
Unprotected vs protected groundwater	3.5 (0.7, 18.5)	0.14	1.2 (0.2, 6.1)	0.84
Other MDG regions vs sub-Saharan Africa	1.4 (0.4, 4.8)	0.59	1.2 (0.4, 3.4)	0.75
Longitudinal vs cross sectional	3.9 (0.8, 19.4)	0.09	2.4 (0.6, 9.8)	0.24
High quality vs low quality	1.5 (0.4, 5.5)	0.51	1.7 (0.5, 5.5)	0.38
Published in or before 2009 vs after 2009	4.0 (1.4, 11.6)	0.01	3.9 (1.5, 10.4)	0.007
Non-random selection vs random selection	3.9 (0.9, 17.7)	0.07	3.9 (0.7, 21.7)	0.12
FIB: TTC vs E. coli	1.3 (0.4, 4.2)	0.61	1.0 (0.4, 2.9)	0.99

Data were derived from bivariate random effects regression of the number of noncompliant samples out of the total number of samples. All adjusted models included one covariate of interest an interaction with source noncompliance and an interaction term with HSW noncompliance. All levels of a variable were included in the models; however, if a variable had more than two levels, only two were selected to calculate an odds ratio.

Table 5: Odds ratios to comparing source and HSW percent noncompliance.

Contrast	OR (95% CI)	p-value
Source vs HSW (unadjusted)	3.5 (2.5, 5.0)	<0.001
Source vs HSW (adjusted[1])	2.3 (1.4, 3.9)	0.001

[1]Adjusted for water supply type using the categorization from Table 2

Figure Legends

Figure 1: Selection of articles for meta-analysis

Figure 2: % noncompliance of source versus household stored water (HSW) for the 36 studies (55 water supplies) included in a bivariate meta-regression, showing pooled mean % source and HSW noncompliance for piped water versus water from other supplies as estimated through this meta-regression.[1]

[1]Grey lines indicate the confidence intervals of these study data, calculated using the standard formula for standard error of a proportion. The unfilled shapes are mean noncompliance of source and HSW, calculated using a bivariate random effects model with dichotomous source type interacted with both source noncompliance and HSW noncompliance. The 95% confidence regions for mean noncompliance were calculated using the geometric relationships between variance and correlation (Pakula, L 2008). Model estimates and 95% confidence intervals from meta-regression were back transformed to derive estimates of mean compliance and 95% confidence intervals on the scale of the original study data. 95% confidence intervals took the form of a rotated ellipse prior to back transformation.

Figure 3. Funnel plot of log odds ratios of HSW versus source water. Log odds greater than 0 indicates that noncompliance is greater in HSW.

Figure 1.

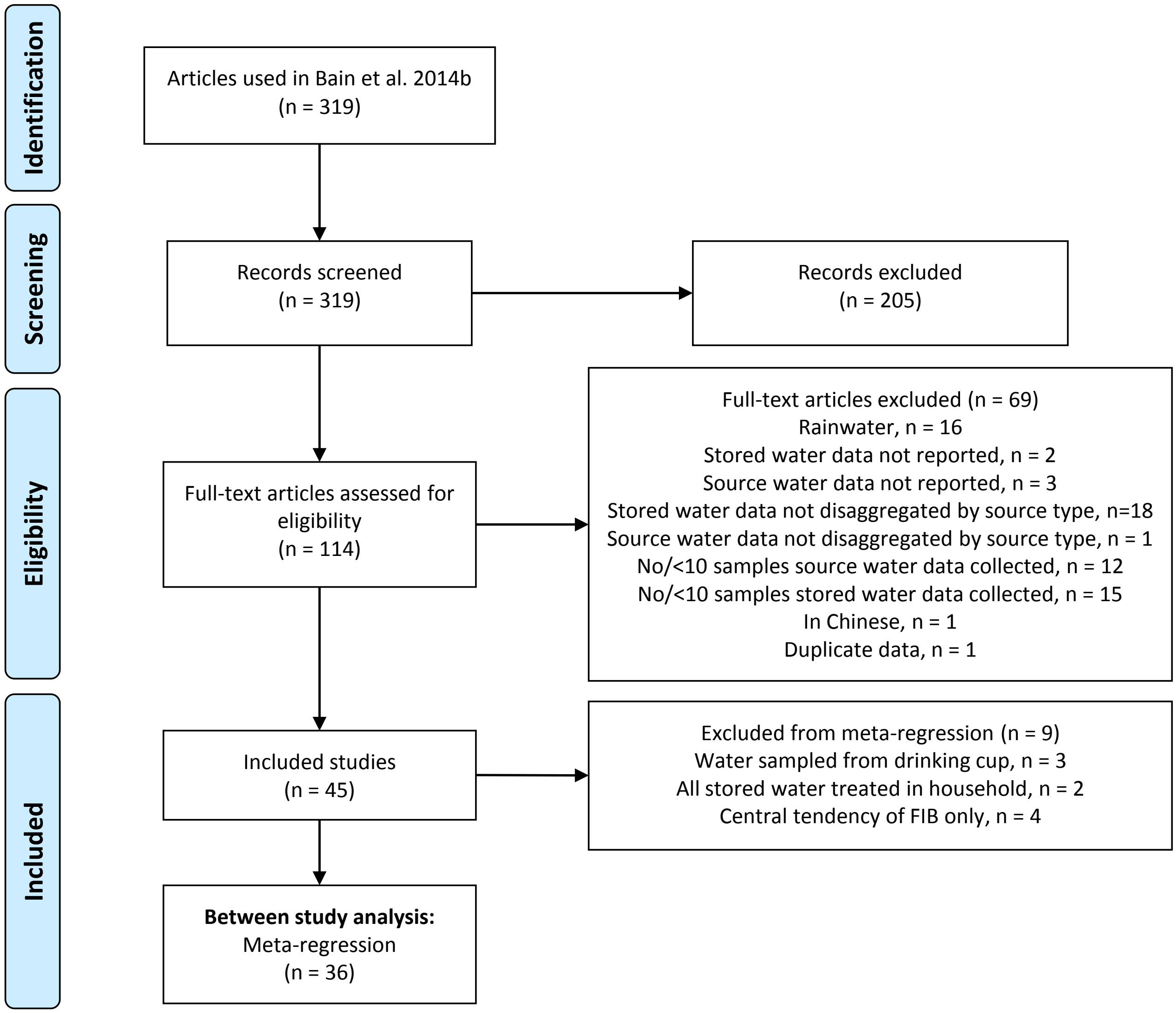

Identification
Screening
Eligibility
Included
Articles used in Bain et al. 2014b
(n = 319)
Records screened
(n = 319)
Records excluded
(n = 205)
Full-text articles assessed for eligibility
(n = 114)
Full-text articles excluded (n = 69)
Rainwater, n = 16
Stored water data not reported, n = 2
Source water data not reported, n = 3
Stored water data not disaggregated by source type, n=18
Source water data not disaggregated by source type, n = 1
No/<10 samples source water data collected, n = 12
No/<10 samples stored water data collected, n = 15
In Chinese, n = 1
Duplicate data, n = 1
Included studies
(n = 45)
Excluded from meta-regression (n = 9)
Water sampled from drinking cup, n = 3
All stored water treated in household, n = 2
Central tendency of FIB only, n = 4
Between study analysis:
Meta-regression
(n = 36)

Figure 2.

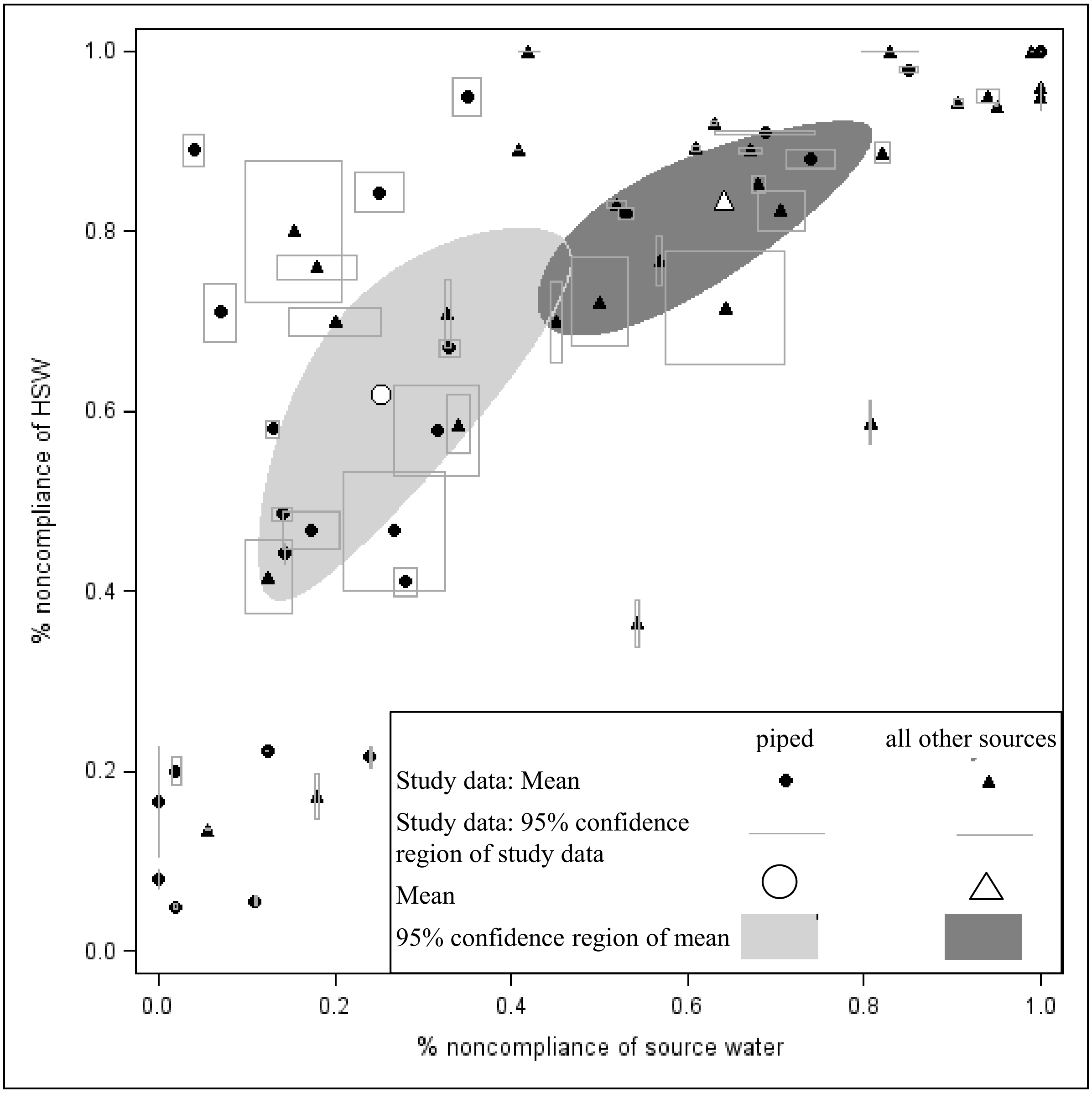
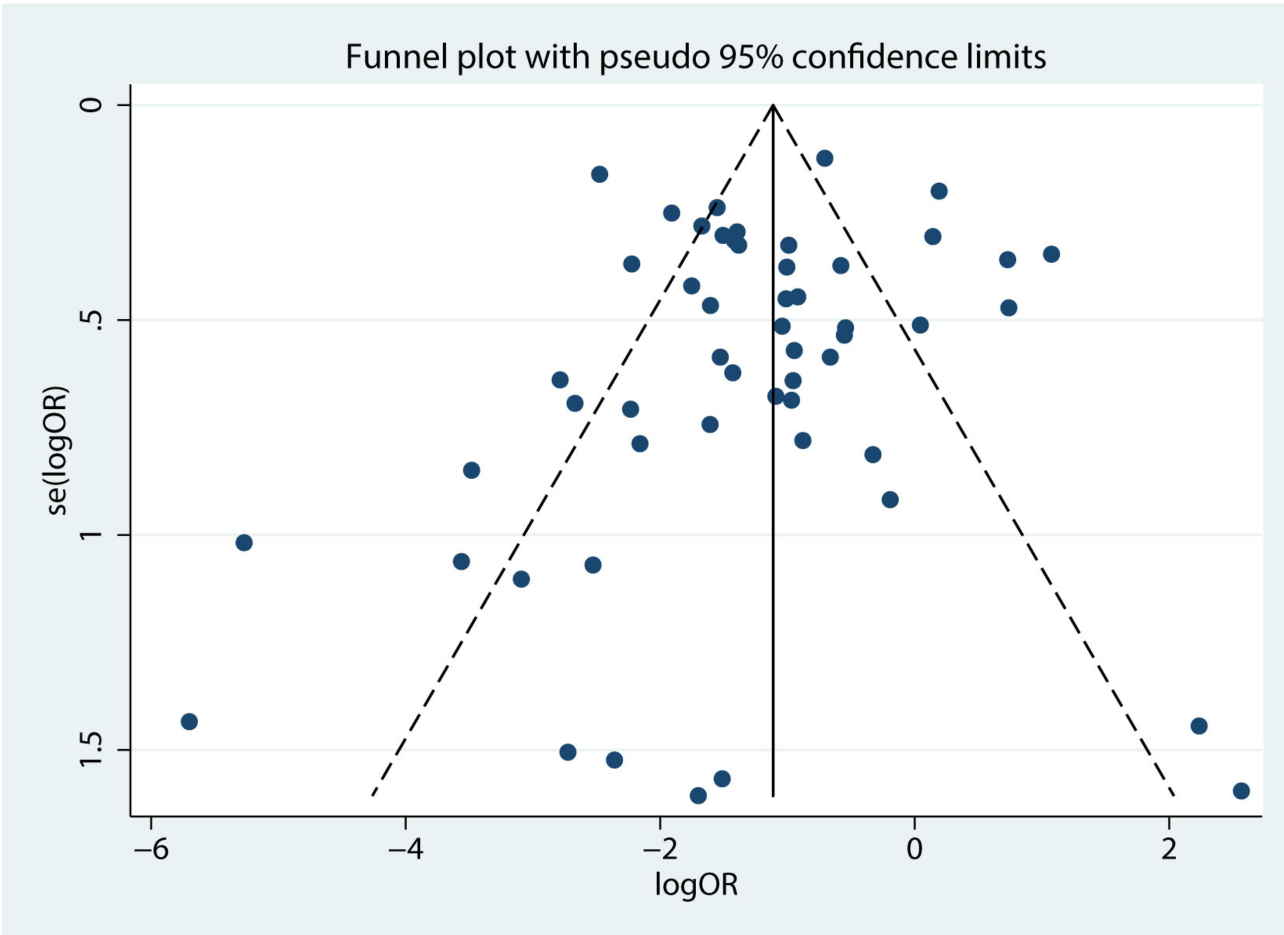

1.0
0.8
0.6
0.4
0.2
0.0
% noncompliance of HSW
0.0
0.2
0.4
0.6
0.8
1.0
% noncompliance of source water
piped
all other sources
Study data: Mean
Study data: 95% confidence
region of study data
Mean
95% confidence region of mean

Figure 3.


Funnel plot with pseudo 95% confidence limits
se(logOR)
0
.5
1
1.5
−6
−4
−2
0
2
logOR