Am. J. Trop. Med. Hyg., 89(2), 2013, pp. 238–245
doi:10.4269/ajtmh.13-0017
Copyright © 2013 by The American Society of Tropical Medicine and Hygiene

A Stepped Wedge, Cluster-Randomized Trial of a Household UV-Disinfection and Safe Storage Drinking Water Intervention in Rural Baja California Sur, Mexico

Joshua S. Gruber,* Fermin Reygadas, Benjamin F. Arnold, Isha Ray, Kara Nelson, and John M. Colford Jr.

University of California, Berkeley, Division of Epidemiology, Berkeley, California; University of California, Berkeley, Energy and Resources Group, Berkeley, California; Fundacion Cantaro Azul, La Paz, Mexico; University of California, Berkeley, Civil and Environmental Engineering, Berkeley, California

Abstract. In collaboration with a local non-profit organization, this study evaluated the expansion of a program that promoted and installed Mesita Azul, an ultraviolet-disinfection system designed to treat household drinking water in rural Mexico. We conducted a 15-month, cluster-randomized stepped wedge trial by randomizing the order in which 24 communities (444 households) received the intervention. We measured primary outcomes (water contamination and diarrhea) during seven household visits. The intervention increased the percentage of households with access to treated and safely stored drinking water (23–62%), and reduced the percentage of households with *Escherichia coli* contaminated drinking water (risk difference (RD): –19% [95% CI: –27%, –14%]). No significant reduction in diarrhea was observed (RD: –0.1% [95% CI: –1.1%, 0.9%]). We conclude that household water quality improvements measured in this study justify future promotion of the Mesita Azul, and that future studies to measure its health impact would be valuable if conducted in populations with higher diarrhea prevalence.

INTRODUCTION

Household water treatment and safe storage (HWTS) interventions are promoted to improve drinking water quality and reduce diarrhea in resource poor settings.[1,2] Several HWTS strategies have been shown to effectively reduce microbial drinking water contamination and diarrhea.[3–5] Much of this evidence, however, is based on efficacy trials, with concentrated follow-up, and there is concern whether HWTS interventions are effective or sustainable outside of intensive research settings.[3,6–8] To identify microbiologically efficacious HWTS interventions that will be effective and sustainable at scale, HWTS studies should incorporate regionally adapted implementation strategies with less intensive research designs.[9]

The objective of this study was to measure the effectiveness of a regional safe drinking water intervention in rural communities of Baja California Sur (BCS), Mexico. The intervention was developed by a local non-profit organization and combined a community-level behavior-change program with installations of a novel ultraviolet (UV)-based HWTS system (the Mesita Azul). We conducted a cluster-randomized, stepped wedge trial to evaluate the impact of the Mesita Azul intervention on drinking water contamination and diarrhea.

METHODS

Study setting. This study took place in rural BCS, Mexico, between August 2009 and January 2011 (Map: Supplemental Appendix A). The study area has an arid climate and consists of dispersed goat and cattle ranching communities with limited access to basic infrastructure, including paved roads, piped water, sewerage, and "grid" electricity; use of small (100 W_p) solar panels is common. Most households rely on wells or springs to collect drinking water, which is typically stored in open containers and extracted with shared dipping utensils.

Fundacion Cantaro Azul (FCA), a non-profit organization based in Mexico, sought this research collaboration to rigorously evaluate the scale up of a previously piloted intervention to 400–500 households in rural BCS. As part of this collaboration FCA agreed to randomize the order in which new communities received the intervention.

Study population. Fundacion Cantaro Azul identified communities where the Mesita Azul would be an appropriate disinfection strategy before this study; final study eligibility was determined during baseline data collection. Eligible clusters (communities): 1) were within La Paz and Los Cabos *municipios* (counties); 2) had not participated in an FCA pilot program; 3) lacked access to centrally treated drinking water, and; 4) collected drinking water from local sources year-round. Within clusters, household samples were tested for arsenic (not removed by the Mesita Azul) and confirmed to be below Mexican limits (< 25 μg/L)[10]; clusters were not included if arsenic was common in baseline drinking water samples. In addition, clusters were ineligible if household water samples commonly had a measured absorption coefficient greater than $0.1 cm^{-1}$ at 254 nm (the wavelength at which UV light is emitted by the lamp used in the Mesita Azul); absorbance levels were also tested and monitored throughout the study.

All full-time residences in a cluster were considered eligible and identified for enrollment at baseline; eligible households that were not available at baseline were approached for enrollment during the first post-baseline visit (Step 1).

Intervention. Fundacion Cantaro Azul designed the Mesita Azul, a HWTS system that incorporates UV Tube disinfection technology (Supplemental Appendix B).[11] Fundacion Cantaro Azul promoted the Mesita Azul through community meetings organized with local leaders followed by household visits to install the systems. During the study period, FCA rolled out two program variations ("Basic" and "Enhanced"). In this article we treat the program variations as a single intervention; however, we explore heterogeneity of results between program variations (Supplemental Appendix B).

Study design. A traditional parallel-arm trial would have required FCA to significantly increase implementation staff, and thereby modify the existing implementation strategy we sought to evaluate (phased rollout). Instead, to avoid significant changes to the existing program, we randomized the order in which communities received the intervention using a stepped

* Address correspondence to Joshua S. Gruber, University of California, Berkeley, Division of Epidemiology, 101 Haviland Hall, Berkeley, CA 94720-7358. E-mail: jsgruber@berkeley.edu

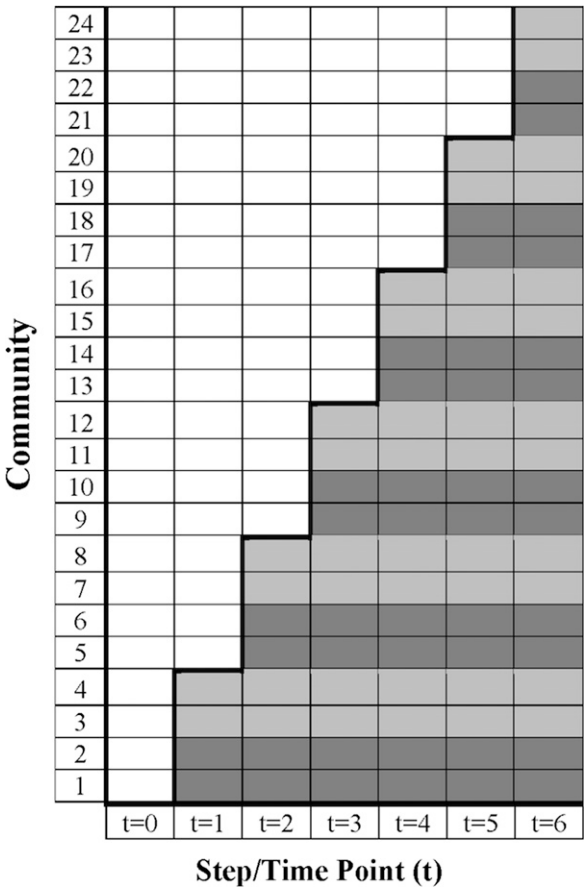

FIGURE 1. Stepped Wedge Schematic for the Mesita Azul Intervention Study. Twenty-four clusters were enrolled at baseline (t = 0) and randomly ordered. All communities started in the control group (white squares). The first four randomly ordered communities (crossover group 1) received the intervention (gray squares) at Step-1 (t = 1). The next four communities (crossover group 2) received the intervention in Step-2, and so on. Within each crossover group two communities were randomized to the Basic Program (dark gray squares) and two to the Enhanced Program (light gray squares). Once a community crossed-over it remained in the intervention group for the remainder of the study. Randomized rollout balances covariates between control- (white squares) and intervention- (all gray squares) periods and creates two comparison groups.

wedge design.[12–14] At baseline, all clusters started in the control arm and, in a random sequence, "crossed over" to the intervention arm during one of six post-baseline "steps" (Figure 1). Post-baseline steps lasted 2 months each. Surveillance was conducted in all study households to measure outcomes at each step. A minimum of 14 days elapsed between Mesita Azul installations in a community and subsequent household surveillance.

Randomization. We used a two-stage randomization process to determine 1) the sequence communities that received the intervention; 2) which communities received the "Basic" and "Enhanced" program variations. At the completion of baseline data collection each eligible community was assigned a random number between zero and one by an investigator (STATA, version 10, College Station, TX). Communities were sequenced based on their random number and organized into crossover groups of four; the first crossover group received the intervention in Step 1, the second in Step 2, etc. (Figure 1). In the second stage, while still in random order, clusters were assigned to one of two groups in an alternating pattern; a coin flip determined who received the Basic program, and the other received the Enhanced program.

Randomization is used to reduce bias and balance community covariates at baseline. In a stepped wedge design, such as the current study, communities contribute different amounts of time to intervention and control periods, making traditional measures of covariate balance between intervention arms difficult to use.[12,15,16] In this study, we used a novel method to assess covariate balance by calculating a weighted average of each baseline characteristic for control and intervention periods; community characteristics were weighted by the amount of person-time they contributed to control and intervention periods (i.e., a cluster that crossed over in Step 2 contributed baseline covariates to two control periods (Steps 0–1) and five intervention periods (Steps 2–6)).

Data collection. We collected baseline and follow-up surveillance data using household questionnaires that we piloted in non-study communities of BCS. A supervisor reviewed completed questionnaires weekly and intermittently observed interviews to ensure consistency. The data collection team was separate and operated independently from implementation activities.

Acquisition, access, and exclusive use. We report acquisition, access to, and exclusive use of household drinking water treatment and safe storage practices as general measures of compliance. The percentage of households that acquired a Mesita Azul during the study period was measured by the presence of a Mesita Azul in the household during any visit, whether or not it was in use. The presence of Mesita Azul is also reported by a visit. A household was considered to have access to treated and safely stored drinking water if any source during a visit was reported to be 1) commercially purified; or 2) disinfected in the home using Mesita Azul, boiling or chlorination, and stored in a narrow mouth container. A household was considered to be an exclusive user (fully compliant at that visit with the messages delivered through the community intervention) if all identified sources of drinking water were reported to be purified or disinfected and were observed to be safely stored.

For the previous measures, reported use of the Mesita Azul or purified water was confirmed by visual inspection and probing by the field team. If a household reported use of the Mesita Azul to disinfect water the field team visually inspected the Mesita Azul for signs of use (e.g., absence of dust, wetness, unit was plugged in). Similarly, purified water is almost exclusively purchased in standard 20-L bottles. If a household reported a source of water to be purified, but the water was not stored in a 20-L bottle, the field team probed to confirm the presence of purchased bottles in the household. Chlorine use was effectively absent from our study at baseline, and plans to confirm chlorine usage using test strips were not implemented.

Water quality. Our primary water quality outcome was the percent of households with contaminated drinking water, classified by *Escherichia coli* ≥ 1 most probable number (MPN)/100 mL, enumerated using Colilert 18 h Quantitray (IDEXX, Westbrook ME). During each visit, drinking water containers were identified in two ways: the respondent was asked to identify the container from which they most recently drank and then to identify which storage containers were the most common sources for themselves, an adult of the opposite gender, and any children; the interviewer prompted the

respondent regarding any visible containers that were not initially identified. Once all "last-glass" and "most-common" sources were identified the respondent was asked to collect water from each as though getting a drink (i.e., with a glass or cup), from which a sample was poured into a 100-mL sterile container. Samples were immediately stored in coolers with ice, and processed within 10 h of collection. Processed samples were incubated at 35°C for 18 h and the number of positive *E. coli* wells was counted under UV light in accordance with IDEXX procedures. Households with ≥ 1 MPN *E. coli*/100 mL in any storage container were categorized as having contaminated drinking water.

Diarrhea. Our primary health outcome was the 7-day prevalence of diarrhea, defined as three or more loose stools in a 24-h period, or one stool with blood or mucus.[17,18] We asked all household members ≥ 15 years of age directly about diarrhea symptoms experienced over the previous 7 days. If a household member was not available we asked another adult with the best knowledge. We asked the closest caretaker about symptoms experienced by members < 15 years of age.

Statistical analyses. Our primary parameter of interest was the average effect of the intervention, which we estimated by intention-to-treat analyses; all households in an intervention cluster were classified as treated, whether they actually acquired or used a Mesita Azul. We compared group means (percent of households with contaminated water and 7-day prevalence of diarrhea) from intervention periods to control periods on absolute (risk difference [RD]) and relative (risk ratio [RR]) scales.

We modeled the outcomes (Y) of interest (the percent of households with contaminated drinking water; Y_{jkt}, and the prevalence of diarrhea; Y_{ijkt}) as

$$E(Y) = \alpha + \beta(A_{kt}) + \sum_{t=1}^{6}\gamma_t I(T=t) + W_j \qquad (1)$$

where $Y_{jkt} = 1$ if household j, in community k at time t had ≥ 1 MPN *E. coli*/100 mL (Y_{jkt} =0 otherwise), and $Y_{ijkt} = 1$ if individual i, in household j, in community k at time t had an episode of diarrhea in the previous 7 days; $A_{kt} = 1$ if community k at time t had the intervention ($A_{kt} = 0$ otherwise), $I(T = t)$ is an indicator of the time-step ($t \in \{0,1,2,3,4,5,6\}$) and γ_t is a fixed-effect of step t on contaminated drinking water or diarrhea. We included the fixed-effect for step to avoid conflating the effect of time with the effect of the intervention[14]; we provide fully unadjusted results (no fixed effects) for comparison purposes. We repeated these analyses restricting the study population to children < 5 years of age at baseline to estimate impacts of the intervention among this sub-population. In secondary analyses, we added an additional set of baseline covariates ($\boldsymbol{W_j}$) to the model to account for any imbalances in baseline covariates, or differential missingness between treatments and control time periods. In bivariate analyses, covariates that were associated with contaminated drinking water (P value on the Wald statistic ≤ 0.20) were added to model (1) above, in a forward stepwise fashion, starting with the covariates with the lowest P values. Covariates that changed the estimate of the effect of treatment (β) by > 5% were included in $\boldsymbol{W_j}$ in the final adjusted model. Covariates ($\boldsymbol{W_{ij}}$) were selected in the same manner for diarrhea except that models with a full set of identified covariates did not converge because of sparse data (rare outcome). Instead, a reduced set of covariates was selected (covariates with the smallest P values were added to the model and included if the coefficient on treatment changed by > 5%; if the model did not converge that covariate was skipped until all covariates were tested.).

Outcomes were modeled on both linear and log-Poisson scales. On the linear scale, β was estimated using ordinary least squares regression and interpreted as the average difference in the percent of households with contaminated drinking water, or average difference in prevalence of diarrhea (RD) in intervention time periods compared with control time periods after controlling for step. Confidence intervals were constructed using robust standard errors to account for clustering at the community (cluster) level. On the log-Poisson scale, β estimates the average relative probability (RR) of having the outcome of interest comparing intervention periods to control periods, after controlling for step. Log-Poisson regressions were estimated using generalized estimating equations and robust SEs clustered at the community level.[19]

In separate analyses, the effect of each program (Basic/Enhanced) on water quality and diarrhea was estimated by replacing the single treatment indicator (A_{kt}) with indicators of program assignment (B_{kt}; Basic) and (E_{kt}; Enhanced):

$$E(Y) = \alpha + \beta_1(B_{kt}) + \beta_2(E_{kt}) + \sum_{t=1}^{6}\gamma_t I(T=t). \qquad (2)$$

In model (2), the coefficients β_1 and β_2 estimate the effect of each respective program on the outcome of interest (contaminated water or diarrhea) compared with non-intervention clusters, on both linear and log-Poisson scales, while controlling for step.

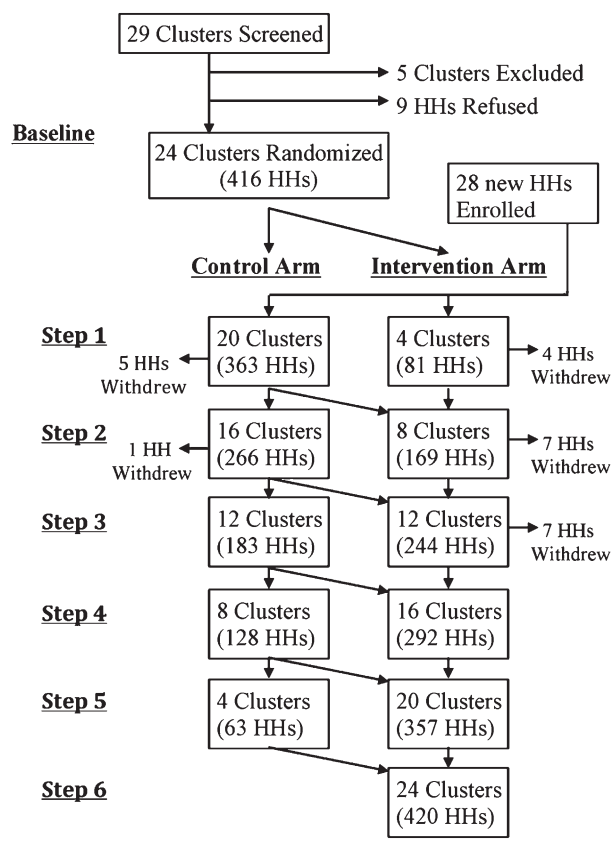

FIGURE 2. Flow Diagram of the Mesita Azul Trail Study Population. After baseline, four households (HHs) crossed over to the intervention in each step.

Power calculations. Power calculations were conducted in R (http://www.r-project.org/) using a previously published approach.[14] Funding available to FCA to supplement system and implementation costs limited the number of households that could be included in the trial to 400–500. *A priori* it was estimated that the study would have > 80% power to detect 12–17% relative reductions in the percent of households with contaminated drinking water if between 350 and 550 households were enrolled, assuming a 50% baseline prevalence of contaminated drinking water and an intra-cluster correlation coefficient (ICC) of 0.1. Assuming an average of 1,530 observations per step (4.5 individuals per household, in 400 households and an average follow-up rate of 85%) over seven steps, and an ICC = 0.1, we estimated *a priori* that this study would have > 80% power to detect a 23% relative reduction in diarrhea if prevalence in control periods was 10%; 35% relative reduction if control period prevalence was 5%; and a 50% relative reduction if control period prevalence was 2.5%.

Ethical considerations. The Office for the Protection of Human Subjects at the University of California, Berkeley approved all research procedures and all households provided informed consent.

Trial registration. This study was registered at ClinicalTrials.gov (NCT01637389).

RESULTS

Study population. We excluded five clusters at baseline that did not meet eligibility criteria. The remaining 24 clusters were randomly ordered for rollout of the intervention (Figure 2). Within eligible clusters, we enrolled 444 (92%) eligible households overall (1915 individuals): 28 households (121 individuals) were unavailable during baseline and were enrolled during Step 1; 184 new individuals moved into study households after enrollment. Only 110 children under 5 years of age were enrolled in the study. Eligible households were not enrolled for various reasons: 9 (2%) declined to participate; 10 (2%) were not identified at baseline; 14 (3%) could not be contacted; and 5 (1%) other reasons.

Characteristics were well balanced between intervention and control periods (Table 1). By chance, larger clusters (more households) crossed over earlier and contributed more person-time overall to intervention periods (6,477 versus 5,640 person-weeks); this did not create imbalance between intervention and control periods in other measurable characteristics.

Missing surveillance. The 444 enrolled households could be visited a maximum of seven times (3,108 possible observations). On average, we collected information on 85% of individuals and 84% of households at each step (Supplemental Tables 1

TABLE 1

Comparison of baseline characteristics weighted by time contributed to control periods and intervention periods*

Characteristic	Baseline	Control periods	Intervention periods
All Ages	*N* = 1,731	5,640 person-weeks	6,477 person-weeks
Age in years, mean (sd)	38 (23.8)	38 (23.8)	37 (23.8)
Demographics, *n* (%)			
Female	790 (46%)	2,524 (45%)	3,006 (46%)
Under 15 years	349 (20%)	1,071 (19%)	1,372 (21%)
Under 5 years	110 (6%)	354 (6%)	415 (6%)
Baseline illness, *n* (%)†			
Diarrhea‡	79 (5%)	233 (4%)	320 (6%)
Adults (> 15 years)	*N* = 1,382	4,569 person-weeks	5,105 person-weeks
Employment and education, *n* (%)			
Traditional ranching activities§	1,020 (75%)	3427 (76%)	3,713 (74%)
Not working§	25 (2%)	68 (1%)	107 (2%)
No education¶	278 (20%)	930 (20%)	1,016 (20%)
Households	*N = 444*	*N = 1470*	*N = 1638*
Water quality, *n* (%)			
< 1 MPN *E. coli*/100 mL‖	158 (40%)	531 (39%)	575 (41%)
< 10 MPN *E. coli*/100 mL‖	271 (69%)	909 (68%)	988 (70%)
Hygiene and sanitation, *n* (%)			
Bar soap at wash station**	216 (54%)	725 (53%)	787 (55%)
Feces in Yard (human or animal)††	143 (35%)	517 (36%)	484 (33%)
Adequate sanitation‡‡l	216 (49%)	688 (47%)	824 (50%)
Self-reported open defecation‡‡	43 (10%)	179 (12%)	122 (7%)
Programs targeting the poor *n* (%)			
Seguro popular insurance‡‡	266 (60%)	870 (59%)	991 (61%)
Oportunidades§§	284 (64%)	929 (64%)	1,058 (65%)
Infrastructure, *n* (%)			
Live on improved road¶¶	102 (23%)	351 (24%)	363 (23%)
Palm roof¶¶	149 (34%)	472 (32%)	570 (35%)
Dirt floor¶¶	322 (73%)	1,014 (69%)	1,240 (76%)
HH possessions, *n* (%)			
Functional radio‡‡	300 (68%)	1,022 (70%)	1,078 (66%)
Functional TV‡‡	194 (44%)	619 (42%)	739 (45%)
Functional car‡‡	352 (79%)	1,176 (80%)	1,288 (79%)
Wood burning stove only‡‡	163 (37%)	583 (40%)	558 (34%)
Functional refrigerator‡‡	75 (17%)	212 (14%)	313 (19%)

*See Methods for description of weighting and group definitions; †7-day prevalence; Missing observations at baseline: ‡*N* = 147; §*N* = 14; ¶*N* = 7; ‖*N* = 53; *****N* = 43; ††*N* = 30; ‡‡*N* = 1; §§*N* = 2; ¶¶*N* = 2; 1Adequate sanitation: sealed pit latrine or flush system.

TABLE 2
Household access to and exclusive use of treated and safely stored water*

Safe water strategy	Access and use frequencies (%)			Unadjusted estimate†
	Baseline (N = 415)	Control periods (N = 1,255)	Intervention periods (N = 1,346)	Risk difference (95% CI)
Mesita Azul present‡	2 (0%)	6 (0%)	877 (65%)	65% (57%, 71%)
Access to, n (%)§				
Mesita Azul	0 (0%)	1 (0%)	691 (51%)	51% (45%, 58%)
Commercial purification	84 (20%)	282 (22%)	163 (12%)	–10% (–17%, –4%)
Boiling	1 (0%)	6 (0%)	1 (0%)	–
Chlorine	1 (0%)	2 (0%)	1 (0%)	–
Any of above	86 (21%)	289 (23%)	839 (62%)	39% (29%, 49%)
Exclusive use, n (%)¶				
Mesita Azul	0 (0%)	1 (0%)	533 (40%)	40% (35%, 44%)
Commercial purification	79 (19%)	233 (19%)	131 (10%)	–9% (–15%, –3%)
Boiling	1 (0%)	2 (0%)	0 (0%)	–
Chlorine	1 (0%)	2 (0%)	1 (0%)	–
Any of above	81 (20%)	240 (19%)	679 (50%)	31% (23%, 40%)

*Commercial purification, boiling, and chlorination were the only safe water strategies observed at baseline.
†Reported risk differences are unadjusted; step-adjusted results were slightly larger in magnitude but equivalent.
‡During a visit, a Mesita Azul was observed in the household (whether in use or not); two households acquired a Mesita Azul from elsewhere in BCS before the study.
§During a visit, any source of drinking water was treated by a given strategy and stored in a narrow mouth container.
¶During a visit, all sources of drinking water were treated by a given strategy and stored in a narrow mouth container.

and 2): 73% of households missed no more than one interview; 6.5% missed five or six interviews. Missing observations were more likely to have had low-risk drinking water (< 10 *E. coli*) at baseline; otherwise they were not systematically different from households that contributed to follow-up surveillance (Supplemental Tables 1 and 2).

Acquisition, access, and exclusive use. Among the study population, 302 households (68%) acquired a Mesita Azul. A Mesita Azul was present in 65% of intervention visits; exclusive use of Mesita Azul treated and safely stored water was observed in 40% of intervention visits (Table 2). At baseline, 21% of households had treated and safely stored water—commercial purification was the most common strategy (20%; Table 2). Exclusive use of purified water decreased during intervention visits (19–10%; Table 2). However, overall access to (23–62%) and exclusive use of treated and safely stored water (19–50%) increased during intervention visits (Table 2).

Water quality. At baseline, 60% [SE: 2.0%] of households had detectable *E. coli* in at least one of their drinking water containers. During Step 1, the percent of households in intervention clusters with detectable *E. coli* dropped to 30%, and remained lower than control clusters throughout the study, despite seasonal variations (Figure 3).

During intervention periods, 43% [SE: 1.4%] of households had detectable *E. coli*, compared with 59% [SE: 1.4%] during control periods (Supplemental Table 3). After controlling for step (to avoid conflating an effect of time with the effect of the intervention), we observed a significant reduction in the percentage of households in intervention periods with detectable *E. coli* (RD: –19% [95% CI: –0.26%, –0.12%]; Table 3). Controlling for additional covariates and step (RD: –15% [95% CI: –21%, –10%]), and restricting the data to households enrolled at baseline (RD: –18% [95% CI: –25%, –11%]), did not change our inference (Table 3).

Diarrhea. At baseline, the overall 7-day prevalence of diarrhea was 5.0% [SE: 0.5%]. During Step 1, diarrhea prevalence dropped in both intervention and control clusters and remained between 1.2% and 3.5% for the remainder of the study (Figure 4).

The average 7-day prevalence of diarrhea was 3.1% [SE: 0.2%] during control periods and 2.3% [SE 0.2%] during intervention periods (Supplemental Table 3). After controlling for step, control and intervention periods did not differ at the 95% confidence level (RD: –0.1% [95% CI: –1.2%, 0.9%]; RR: 0.80 [95% CI: 0.51, 1.27]; Table 3). The results did not differ after controlling for additional covariates and restricting data to individuals enrolled during baseline (Table 3).

Among children < 5 years of age diarrhea prevalence was 2.7% [SE: 0.8%] during control periods, and 1.2% [0.5%] during intervention periods. After controlling for step there was no difference in diarrhea prevalence (RD: 0.0% [95% CI: –1.5%, 1.6%]); models estimating RR did not converge as a result of sparse data (Table 3).

Program variations. The Basic and Enhanced program variations had similar impacts on contaminated drinking water (Supplemental Figure 1) and diarrhea (Supplemental Figure 2) when compared with control periods and to each other (Supplemental Table 4).

DISCUSSION

Summary. In this study, we found evidence that the Mesita Azul safe drinking water intervention significantly reduced

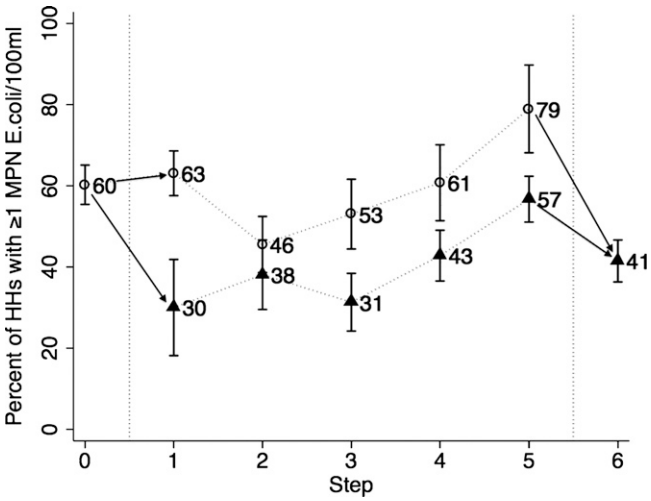

FIGURE 3. Percent of households with contaminated drinking water by treatment arm and step (black triangles: intervention households; open circles: control households; black bars: 95% confidence intervals).

TABLE 3
Effect of the Mesita Azul intervention on drinking water contamination and diarrhea

Analysis	*N*	Outcome frequency-control periods (%)	Risk difference (95% CI)	Risk ratio (95% CI)
Water contamination				
ITT_{step}*	2,436	698 (59%)	−19% (−26%, −12%)	0.67 (0.58, 0.77)
$ITT_{covariates}$†	2,297	690 (59%)	−15% (−21%, −10%)	0.72 (0.64, 0.80)
$ITT_{sub\text{-}pop}$*‡	2,309	694 (59%)	−18% (−25%, −11%)	0.70 (0.63, 0.78)
Diarrhea				
ITT_{step}*	10,854	157 (3.1%)	−0.1% (−1.1%, 0.9%)	0.80 (0.51, 1.27)
$ITT_{covariates}$§	10,779	157 (3.1%)	−0.1% (−1.1%, 0.8%)	0.89 (0.57, 1.38)
$ITT_{sub\text{-}pop}$*‡	10,244	153 (3.1%)	−0.3% (−1.3%, 0.7%)	0.79 (0.49, 1.29)
Diarrhea: < 5 years				
ITT_{step}*	765	9 (2.7%)	0.0% (−1.5%, 1.6%)	DNC

*Model adjusted for step (season).
†Model adjusted for step and baseline covariates: presence of feces in yard, ownership of a functional refrigerator.
‡Restricted to households enrolled at baseline.
§Model adjusted for step and baseline covariates: household reported personal hygiene prevents diarrhea at baseline, head of household started primary school.
ITT = intention-to-treat; CI = confidence interval; DNC = did not converge.

the percentage of households with detectable *E. coli* (−19% [95% CI: −26%, −12%]; Table 3). Yet, we found no detectable reduction in diarrhea (RD: −0.1% [95% CI: −1.1%, 0.9%]); Table 3). The observed relative reduction in diarrhea (RR: 0.80 [95% CI: 0.51, 1.26]) was consistent in magnitude and direction with previously published HWTS studies,[4] but we cannot rule out the possibility that reductions resulted from chance variation. We similarly found no effect of the intervention on diarrhea among children < 5 years of age (RD: 0.0% [95% CI: −1.5%, 1.6%]). However, the "under-five" sample was very small (N = 110). Among enrolled households, 68% acquired a Mesita Azul during the study. We observed overall increases in access to and exclusive use of treated and safely stored water (Table 2), but unsafe water practices were present in 50% of intervention visits.

To our knowledge, this is the first study to evaluate the health impacts of a UV-based HWTS technology in a low- or middle-income country. Additionally, although one previous study reported using elements of a stepped wedge design for a water and sanitation intervention,[19] this is the first HWTS study to be analyzed as a *randomized* stepped wedge trial.

Interpretation. We hypothesized that the intervention would reduce the transmission of waterborne pathogens through drinking water and thereby mediate reductions in diarrhea. Several results suggest that measured reductions in drinking water contamination are attributable to the intervention, rather than to bias: successful randomization (Table 1), infrequent and non-systematically missing data (Supplemental Tables 1 and 2), and an objective outcome measure (MPN *E. coli*/100 mL). Furthermore, classifying households as "contaminated" if *any* storage container had detectable *E. coli* means that our water quality effect estimates are likely conservative. Despite improvements in drinking water quality there are several possible explanations of why we did not observe reductions in diarrhea.

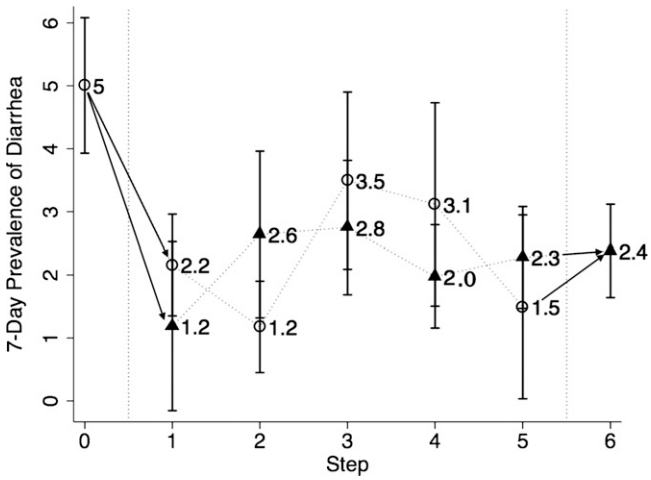

FIGURE 4. Seven-day prevalence of diarrhea by treatment arm and step (black triangles: intervention households; open circles: control households; black bars: 95% confidence intervals).

First, baseline diarrhea prevalence was 5.0% but subsequently fell to 1.2–3.5% for the remainder of the study (Figure 4), which was at the low end of the range of diarrhea prevalence considered for our *a priori* power calculations. Low diarrhea prevalence in control clusters reduced our power to detect a difference in diarrhea prevalence attributable to the intervention. Increasing the number of surveillance visits could have increased the precision of our estimates. However, our goal was to use a research design that minimized the intensity of evaluation activities, and it is possible that increasing surveillance visits would have simultaneously exacerbated reporting biases.[20]

Second, there are two possible mechanisms by which our water quality measures may not have fully captured the underlying risk of diarrhea from drinking water during the study. Indicator bacteria (*E. coli*) are imperfect measures of pathogen exposure[21]; although our results provide strong evidence that the intervention increased disinfection practices, our measures of *E. coli* could have overestimated the presence of diarrheagenic pathogens and thus health benefits of disinfection. Furthermore, modeling efforts suggest that consuming contaminated water even a small fraction of the time can attenuate diarrhea reductions attributable to household water treatment.[22,23] Although the intervention increased treatment and safe storage practices overall, we still observed the presence of untreated or improperly stored drinking water during 50% of intervention visits (Table 2); less than perfect compliance could have diminished the overall health benefits of increased disinfection practices within the household. In addition, we did not attempt to measure the consumption of unsafe water outside of the household, and participant exposure to untreated water while away from home could have also reduced the health impact of increased household treatment and safe storage practices.

Finally, our intervention targeted household drinking water. In settings where sanitation and/or hygiene are important

transmission routes for diarrheal illness we would expect to see smaller impacts from a drinking water intervention.[24] However, no sanitation or hygiene indicator was independently associated with diarrhea, with the possible exception of garbage in the yard or house (Supplemental Table 5).

Limitations. The nature of the intervention prevented us from blinding participants or enumerators; use of subjective outcomes (diarrhea) in an unblinded trial could lead to biased results.[12,25] Indeed, the sharp drop-off in reported diarrhea symptoms between baseline and follow-up in our study could indicate a reporting bias.[7] However, this secular reduction in diarrhea was non-differential by intervention arm (Figure 4), which could have biased our effect estimates towards the null, and would be consistent with a finding of no-effect.[7,26] Future studies could complement self-reported diarrhea with objective outcomes, such as the collection of salivary, blood, or stool specimens to test for pathogen-specific antibody responses.[27] Most HWTS interventions target children < 5 years of age, as this population globally bears the heaviest burden of diarrhea morbidity and mortality.[28,29] Our study included only 110 children < 5 years of age, which limited our ability to estimate effects among this subpopulation. Future studies in populations with higher disease burden, and a larger under-five sub-population would be beneficial.

Use of the stepped wedge design for HWTS trials. To our knowledge, this is the first study to use a randomized stepped wedge design to evaluate the water quality and health impacts of an HWTS intervention. The stepped wedge design allowed us to conduct this rigorous, randomized study without changing the intervention for evaluation purposes, and could be valuable for future evaluations of water, sanitation, or hygiene interventions that are being rolled out or phased-in by regional implementers. A limitation of the stepped wedge design, compared with parallel-arm designs, is that it conflates time effects with intervention effects in situations where calendar time (or time on surveillance) affects the outcome.[14] In our study, calendar time impacted water quality (Figure 3) and reported diarrhea prevalence (Figure 4). However, we avoided conflating the effect of time with the intervention effect by including a fixed-effect for step[14] (unadjusted results: Supplemental Table 3; step-adjusted results: Table 3). When implementation or ethical conditions make parallel-arm trials impossible, the stepped wedge design can be a useful alternative as long as investigators account for the effect of time in the design and analysis stage.[12–14] As an additional contribution to the literature, we show a parsimonious and intuitive method to evaluate covariate balance between study arms in a stepped wedge trial (Table 1).

CONCLUSION

This randomized stepped wedge trial showed that the regional Mesita Azul intervention increased access to and exclusive use of treated and safely stored drinking water and decreased the proportion of households with detectable *E. coli.* We found no statistically significant reduction in diarrhea as a result of the intervention, possibly caused by the consumption of untreated water, or unexpectedly low diarrhea prevalence in control clusters that allowed little room for improvement. The water quality results strongly support the distribution of the Mesita Azul through a safe water program. Given the large improvements in water quality shown in this study, future studies to measure the effect of the Mesita Azul on diarrhea would be valuable, but would require a larger study or a study population with higher baseline diarrhea prevalence.

Received January 10, 2013. Accepted for publication March 20, 2013.

Published online June 3, 2013.

Note: Supplemental materials appear at www.ajtmh.org.

Acknowledgments: We acknowledge the contributions of Ian Balam, the Cantaro Azul field teams, and the Mexican National Council on Science and Technology (CONACYT) for its indirect support through its Postgraduate Studies Scholarships Program.

Financial support: This work was supported by the Sustainable Products & Solutions Program and the Blum Center for Developing Economies, both at the University of California, Berkeley.

Disclosure: Fermin Reygadas is a PhD candidate at UC Berkeley, and co-founder and current Director General of Fundacion Cantaro Azul.

The funders played no role in the study design, the development of this manuscript, or the decision to publish.

Authors' addresses: Joshua S. Gruber, Benjamin F. Arnold, and John M. Colford, Jr., University of California, Berkeley, Division of Epidemiology, Berkeley, CA, E-mails: jsgruber@berkeley.edu, benarnold@berkeley.edu, and jcolford@ucberkeley.edu. Fermin Reygadas, University of California, Berkeley, Energy and Resources Group, Berkeley, CA, and Fundacion Cantaro Azul, La Paz, Mexico, E-mail: reygadas@berkeley.edu. Isha Ray, University of California, Berkeley, Energy and Resources Group, Berkeley, CA, E-mail: isharay@berkeley.edu. Kara Nelson, University of California, Berkeley, Civil and Environmental Engineering, Berkeley, CA, E-mail: nelson@ce.berkeley.edu.

REFERENCES

1. Lantagne DS, Quick R, Mintz ED, 2006. Household water treatment and safe storage options in developing countries: a review of current implementation practices. *Wilson Quarterly, Woodrow Wilson International Center for Scholars Environmental Change and Security Program*. Available at: http://www.wilsoncenter.org/publication/household-water-treatment-and-safe-storage-options-developing-countries-review-current.
2. 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:* 1565–1570.
3. Arnold BF, Colford JM Jr, 2007. Treating water with chlorine at point-of-use to improve water quality and reduce child diarrhea in developing countries: a systematic review and meta-analysis. *Am J Trop Med Hyg 76:* 354.
4. Clasen T, Schmidt WP, Rabie T, Roberts I, Cairncross S, 2007. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. *BMJ 334:* 782.
5. Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, Colford JM, 2005. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet Infect Dis 5:* 42–52.
6. Hunter PR, 2009. Household water treatment in developing countries: comparing different intervention types using meta-regression. *Environ Sci Technol 43:* 8991–8997.
7. Schmidt WP, Cairncross S, 2009. Household water treatment in poor populations: is there enough evidence for scaling up now? *Environ Sci Technol 43:* 986–992.
8. 3Sobsey MD, Stauber CE, Casanova LM, Brown JM, Elliott MA, 2008. Point of use household drinking water filtration: a practical, effective solution for providing sustained access to safe drinking water in the developing world. *Environ Sci Technol 42:* 4261–4267.
9. Clasen T, Bartram J, Colford J, Luby S, Quick R, Sobsey M, 2009. Comment on household water treatment in poor populations: is there enough evidence for scaling up now? *Environ Sci Technol 43:* 5542–5544.

10. Secretaría de Salud de Mexico, 1994. MODIFICACION a la Norma Oficial Mexicana NOM-127-SSA1-1994, Salud ambiental. Agua para uso y consumo humano. Límites permisibles de calidad y tratamientos a que debe someterse el agua para su potabilización. Mexico: Secretaría de Salud de Mexico.
11. Brownell SA, Chakrabarti AR, Kaser FM, Connelly LG, Peletz RL, Reygadas F, Lang MJ, Kammen DM, Nelson KL, 2008. Assessment of a low-cost, point-of-use, ultraviolet water disinfection technology. *J Water Health 6:* 53–65.
12. Brown CA, Lilford RJ, 2006. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol 6:* 54.
13. Gambia Hepatitis Study Group, 1987. The Gambia hepatitis intervention study. *Cancer Res 47:* 5782–5787.
14. Hussey MA, Hughes JP, 2007. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials 28:* 182–191.
15. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T; CONSORT GROUP (Consolidated Standards of Reporting Trials), 2001. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med 134:* 663–694.
16. Friedman LM, Furberg CD, DeMets DL, 2010. *Fundamentals of Clinical Trials*. New York, NY: Springer.
17. Baqui AH, Black RE, Yunus M, Hoque AR, Chowdhury H, Sack RB, 1991. Methodological issues in diarrhoeal diseases epidemiology: definition of diarrhoeal episodes. *Int J Epidemiol 20:* 1057.
18. Arnold BF, Galiani S, Ram PK, Hubbard AE, Gertler PJ, Colford JM, 2013. Optimal recall period for caregiver-reported illness in risk factor and intervention studies: a multicountry study. *Am J Epidemiol 177:* 361–370.
19. Bailey I, Archer L, 2004. The impact of the introduction of treated water on aspects of community health in a rural community in Kwazulu-Natal, South Africa. *Water Sci Technol 50:* 105–110.
20. Zwane AP, Zinman J, Van Dusen E, Pariente W, Null C, Miguel E, Kremer M, Karlan DS, Hornbeck R, Giné X, Duflo E, Devoto F, Crepon B, Banerjee A, 2011. Being surveyed can change later behavior and related parameter estimates. *Proc Natl Acad Sci USA 108:* 1821–1826.
21. Levy K, Nelson KL, Hubbard A, Eisenberg JNS, 2012. Rethinking indicators of microbial drinking water quality for health studies in tropical developing countries: case study in northern coastal Ecuador. *Am J Trop Med Hyg 86:* 499–507.
22. Brown J, Clasen T, 2012. High adherence is necessary to realize health gains from water quality interventions. *PLoS ONE 7:* e36735.
23. Enger KS, Nelson KL, Rose JB, Eisenberg JS, 2012. The joint effects of efficacy and compliance: a study of household water treatment effectiveness. *Water Res 47:* 1181–1190.
24. Eisenberg JN, Scott JC, Porco T, 2007. Integrating disease control strategies: balancing water sanitation and hygiene interventions to reduce diarrheal disease burden. *Am J Public Health 97:* 846.
25. Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, Gluud C, Martin RM, Wood AJ, Sterne JA, 2008. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ 336:* 601–605.
26. Rothman K, Greenland S, 1998. *Modern Epidemiology*. Second edition. Philadelphia, PA: Lippincott-Raven Publishers.
27. Crump JA, Mendoza CE, Priest JW, Glass RI, Monroe SS, Dauphin LA, Bibb WF, Lopez MB, Alvarez M, Mintz ED, Luby SP, 2007. Comparing serologic response against enteric pathogens with reported diarrhea to assess the impact of improved household drinking water quality. *Am J Trop Med Hyg 77:* 136–141.
28. WHO, 2008. *The Global Burden of Disease: 2004 Update*. Geneva: World Health Organization.
29. WHO, 2010. *Diarrhoeal Disease*. Available at: http://www.who.int/mediacentre/factsheets/fs330/en/index.html. Accessed June 15, 2010.