PEER REVIEW

Assessing operational performance benefits of a Water Safety Plan implemented in Southwestern France

Authors

K Setty iD
*University of North Carolina at Chapel Hill, Department of Environmental Sciences and Engineering, 166 Rosenau Hall, CB #7431Chapel Hill, NC 27599-7431
Email: ksetty@live.unc.edu*

G O'Flaherty
The Water Institute at University of North Carolina, Chapel Hill, NC, USA

J Enault
Suez, Centre International de Recherche sur l'Eau et l'Environnement (CIRSEE), Paris, France

S Lapouge
Suez Eau France, Bordeaux, France

JF Loret
Suez, Centre International de Recherche sur l'Eau et l'Environnement (CIRSEE), Paris, France

J Bartram
The Water Institute at University of North Carolina, Chapel Hill, NC, USA

Corresponding author:
Karen Setty, as above

Abstract

Aims: The World Health Organization (WHO) has recommended Water Safety Plans (WSPs) since 2004 as a means to reduce drinking water contamination and risks to human health. These risk management programs have shown promise across several potential areas of evaluation, such as economic benefits and regulatory compliance. Since WSPs are largely carried out by people who interact with water treatment equipment and processes, operational performance indicators may be key to understanding the mechanisms behind desirable WSP impacts such as water quality and public health improvement.

Method: This study reports performance measures collected at a WSP implementation location in southwestern France over several years.

Results: Quantitative assessment of performance measures supported qualitative reports from utility managers. Results indicate significantly reduced duration of low-chlorine events at one production facility and a significant decrease in customer complaints related to water quality, manifesting reported improvements in operational performance and the customer service culture.

Conclusion: The findings demonstrate some success stories and potential areas of future performance tracking. Cyclical iteration of the WSP can help to achieve continuous quality improvement. Successfully applied evaluation criteria such as the number of water quality complaints or alarm resolution time might be useful across other locations.

INTRODUCTION

Water Safety Plans (WSPs) are a drinking water risk management approach recommended by the World Health Organization (WHO) since 2004.[1] WSPs have been applied in more than 90 countries and are legally required in many.[2] They function by engaging a team of utility managers and operators in an iterative cycle of risk assessment, establishing controls to manage risks, and verifying whether the approach works or needs revision. In contrast to reactive approaches such as finished water quality monitoring and regulatory reporting, water purveyors who use WSPs seek to comprehensively improve preventive maintenance procedures and critical process steps. WSPs are complex (and often multifaceted) public health interventions that may involve 'soft' discrete interventions, such as staff coordination or documentation of procedures, and/or 'hard' interventions, such as infrastructure upgrades.

Existing WSP evaluation frameworks and performance indicators follow a logic model broadly spanning inputs (e.g. funding and time commitment), activities/outputs (e.g. number of team meetings), outcomes (e.g. operational efficiency or cost savings), and impacts (e.g. water quality or health improvements).[3] Theoretically, the early-stage categories, inputs, and activities, show measurable change more quickly than more distal outcomes and impacts such as public health and socioeconomic changes,[3] although further evaluation is needed to describe causal mechanisms and timing of such changes. Recommended outcome indicators span four categories: institutional, operational, financial, and policy.[4] WSP evaluation models would benefit from clarification of the

relationships among different evaluation indicators and the mechanisms by which determinants (barriers or facilitators) affect outcomes and impacts.[5,6] Furthermore, many jurisdictions lack standardized guidance on recommended WSP performance indicators, which would enable wide-scale comparison.[7]

A number of WSP outcomes have been evaluated in the literature, with the greatest weight of evidence supporting operational, financial, and public health benefits.[8–13] A WHO-sponsored evaluation of water systems employing WSPs in the Asia-Pacific region showed significant changes in operations and maintenance practices (e.g. adoption of standard operating procedures or compliance monitoring plans), the number of water safety–related meetings, water quality testing, consumer satisfaction monitoring, and complaint recording.[12] Common benefits reported by utility managers in a five-country study of WSP implementation at 20 water systems included better hazard control (especially awareness of previously overlooked hazards), as well as improved treatment practices, record keeping, and client and health agency confidence.[13] A study of WSPs evaluated across five different water systems in France and Spain showed fairly consistent improvements in compliance with internal and external water quality benchmarks.[11] Furthermore, changes in health status were detected at multiple WSP implementation sites,[10,11] addressing the ultimate goal of public health protection and improvement. Other WSP progress evaluation indicators have likely been applied by individual utilities in practice, but not yet documented in the literature. The WHO has yet to provide specific guidance on indicators for assessing WSP performance but will likely do so in the future as increasing evidence becomes available.

Monitoring and evaluation of WSPs can be enhanced by tracking frequently overlooked qualitative shifts in organizational culture (e.g. record keeping) alongside more commonly required quantitative measures, such as water quality.[12] In addition to demonstrating tangible improvements attributable to WSPs, these measures

may also aid ongoing quality control efforts (e.g. toward achieving performance goals). Tracking the same indicators consistently over time and over multiple water systems can also enable large-scale comparisons (e.g. meta-analyses) and help to clarify the time frame needed to both achieve and observe results.

In this study, we investigated the potential value of several operational performance indicators used for a WSP at a drinking water utility in southwestern France, seeking to validate and supplement outcomes reported in the literature. We collected approximately six years (2012–2017) of operational performance data and used both qualitative and quantitative analyses to describe trends. The goal of the study was to inform development of consistent, measurable performance indicators for application across multiple water systems.

METHODS
Site description
One location in southwestern France was selected based on availability of operational performance data following earlier nested studies at all WSP implementation locations operated by Suez, which described costs, benefits, water quality, compliance, and health.[6,11,13] The metropolitan drinking water system serves about 740,000 people across 23 towns, drawing from a combination of deep-protected groundwater and shallow groundwater under the influence of surface water. The area-wide system includes 102 groundwater extraction points, 140 drinking water treatment facilities, 50 treated water storage tanks, and approximately 3200 km of distribution pipes. The treatment process typically consists of media filtration (sand and granular-activated carbon (GAC) filtration systems), clarification, and ultraviolet (UV) or chlorine disinfection. Parts of the underground collection, storage, and distribution infrastructure have been in place since 1850. The utility's WSP team began meeting 13 November 2012, and the full system received an ISO 22000 certification for drinking water safety management on 20 December 2013.

Within the full service area of 140 treatment facilities addressed by the WSP team, one treatment facility that serves about 44,000 people (about 6% of the total service population) was chosen to model WSP alarm response dynamics and water quality–related customer complaints. The predominately automated treatment process at this facility consists of pre-oxidation/pre-chlorination with chlorine gas (chlorine dioxide prior to 2012), GAC filtration, UV disinfection, and chlorine disinfection. Ongoing status information (e.g. turbidity, water temperature, UV intensity, free chlorine levels, flow rate) is monitored using automated sensors and wirelessly transmitted to a central control station with 24-h staffing. Operators, who have undergone training with the central WSP team, conduct maintenance on-site at least once a week, including manual verification measurements.

During the WSP implementation process, critical and operational controls were developed primarily to maintain a consistent range of free chlorine levels upon treatment and throughout the distribution network. The production control seeks to continually maintain free chlorine residuals above 0.05 mg/l in produced (finished) water. Levels are monitored at 5-min intervals using online sensors. In the distribution system, the operational limit (general operational goal) is to maintain free chlorine levels below 0.2 mg/l, but not to reduce them as low as 0 mg/l for more than 24 h. The critical limit, at which an alarm is triggered, is a free chlorine level greater than 0.3 mg/l, or as low as 0 mg/l for more than 48 h. An alarm in either production or distribution may lead to additional monitoring checks both at the location of detection and in other affected parts of the network. This might involve sensor verification, manual chlorine testing, and, if needed, isolation of affected water batches.

Data collection and analysis
Data were collected using multiple methods. Results from a standard questionnaire about reported costs, benefits, challenges, and facilitators associated with the WSP were obtained from a previous study that reported

amalgamated results across 20 implementation sites but did not report findings from each site individually.[13] Notes assessing overall WSP experience were obtained from a nested study further evaluating five of the implementation sites, which included a written questionnaire, in-person site visit, and semi-structured focus group discussion regarding water quality, compliance, and health outcomes.[11] Interaction with human subjects (WSP team personnel) was approved by the Institutional Review Board (IRB) at the University of North Carolina at Chapel Hill (study #15-2118). Written details related specifically to operational performance, critical controls, or employee practices were extracted from these sources and, in some cases, translated to English.

Monthly data comprising performance indicators tracked across six categories since onset of the WSP in 2012 were shared by the WSP team lead (categories 1–3 described in Table 1), along with logs of alarms and water quality–specific customer complaints. Complementary indicator categories not shown in Table 1 due to privacy concerns aim to (4) ensure communication about the importance of health risk and good business practices to all stakeholders, (5) monitor the quality and the efficiency of work, and (6) check control measures on water treatment plants as well as distribution networks. Progress toward these objectives is regularly reviewed and the indicators are occasionally revised by the WSP team and utility managers. In addition, raw data from free chlorine water quality sensors at the outlet of one treatment facility were provided by the utility to model low-chlorine event dynamics.

Data on performance indicators that were consistently recorded either monthly or annually over several years (Table 1) were plotted in Excel for visualization of trends. Statistical data analysis was conducted using SAS 9.4 software where possible, given sufficient observations and variability in the data. Before–after comparisons used a t test ($\alpha = .05$). Trends over time in monthly count data were assessed using the GENMOD multiple regression procedure with a Poisson distribution and natural

log link function. To control for seasonal variation, we introduced month as an independent categorical (class) variable in the models.

Free chlorine sensor readings at the single production facility were recorded by the utility at a frequency of approximately 5 min from 2010 to 2017. The data were divided into three periods, 'before' WSP team formation, 'during' WSP implementation, and 'after' WSP certification; for this analysis, the 'during' time period from 13 November 2012 to 20 December 2013 was discarded as a period of transition. Water quality values below the lower detection limit (<0.01 mg/l) were set to zero. Free chlorine data likely affected by temporary plant shutdowns or sensor failures were excluded, although consistent historical records were not available to accurately document these events. Because the data were not flagged at the time of recording, the data cleaning approach was not expected to accurately discern true and false events in all cases but allowed us to simulate actual events by applying a standard procedure for both the before and after data sets. Suspected shutdowns or sensor malfunctions were flagged by identifying events where free chlorine values equal to 0 mg/l (the WSP critical limit) were recorded for more than 24 h and then validating evidence of an external issue with the plant manager. This data cleaning procedure erred on the side of including data not definitively known to be problematic.

To integrate the mixed methods of data collection and interpretation, this study took an exploratory sequential (qual → QUAN) approach.[14] That is, the initial results of the qualitative techniques informed processing of quantitative data to validate reported dynamics. Qualitative results were compiled for a single location, which was not previously disaggregated and analyzed independently in other studies.[6,11,13] The data integration process focused on building and merging, seeking complementarity of measures.[15] The results were reviewed and validated via participatory 'member checks' by three management-level personnel at the water utility studied, permitting

refinements to both the quantitative and qualitative data interpretation.

RESULTS
Qualitative reports from the utility provided guidance and supplemental information to assist the investigation of WSP operational performance changes. Based on unpublished questionnaire results collected by Loret et al.,[13] operational benefits specific to the location under study included the following:

- Better control of production and distribution processes (due to critical control point alarms);
- Improved management responsiveness when alarms identified as critical;
- Process optimization/infrastructure improvement (e.g. for the isolation of water 'batches');
- Improved production/distribution management procedures (e.g. enhanced crisis management);
- Better application of procedures and good practices;
- Better knowledge, understanding, and involvement of staff;
- Better data recording and traceability of events;
- Better handling of consumer complaints.

Analysis of qualitative reports from site visit and focus group discussion notes offered improved understanding of the effect of the WSP on management of the treatment facilities. Indicators recommended by utility managers to evaluate WSP performance included the following: progress toward the action plan, the number of significant hazards being managed, exceedances of critical control point limits, rate of inspections and site monitoring, and reactivity to critical alarms. The WSP mainly sought to balance chlorine levels within a desirable range. Managers felt that values above the distribution system upper limits might add to concerns about disinfection byproducts and their relationship to cancer risk.[16] Local residents also sometimes complained about chlorine odor and taste. From the

Table 1

Partial list[a] of performance indicators tracked by the utility across the full water system since the 2012 onset of the WSP

Objectives		2012	2013	2014	2015	2016	2017
1. Monitor potential pollution sources							
1.1	Progress rate toward completed public utility declarations (DUPs)	Y	Y	Y	Y	Y	N/A[b]
1.2.1	Progress rate of vulnerability studies of sites	Y	Y	N[c]	N[c]	N[c]	N[c]
1.2.2	Progress of research studies on the presence of emerging pollutants	Y	Y	Y	Y	Y	Y
1.3.1	Regulatory compliance rate for source water	Y	Y	Y	Y	Y	Y
1.3.2	Self-monitored compliance rate for source water	Y	Y	Y	Y	Y	Y
2. Monitor drinking water quality at production and storage facilities 24/7							
2.1	Regulatory compliance rate at facility outlet	Y	Y	Y	Y	Y	Y
2.2	Self-monitored compliance rate at facility outlet	Y	Y	Y	Y	Y	Y
2.3	Number of low residual chlorine alarms (<0.05 mg/l) at the facility	N[d]	Y	Y	Y	Y	Y
2.4	Rate of completion of actions following health audits	Y	Y	Y	Y	Y	Y
3. Monitor drinking water quality in the distribution network 24/7							
3.1	Regulatory compliance rate for distribution network	Y	Y	Y	Y	Y	Y
3.2	Self-monitored compliance rate for distribution network	Y	Y	Y	Y	Y	Y
3.3	Number of low residual chlorine alarms from quality sensors	N[d]	N[d]	Y	Y	Y	Y
3.4	Availability of readings from quality sensors (%)	Y	Y	Y	Y	Y	Y
3.5	Number of water quality complaints	Y	Y	Y	Y	Y	Y

WSP: Water Safety Plan.
[a]Objectives 3.6 through 6.3 not shown to maintain privacy of internal practices.
[b]Information not yet published.
[c]All mandatory vulnerability studies carried out prior to 2014.
[d]Alarms defined in 2013.

utility managers' perspective, maintaining lower limits for chlorination was important to avoid microbial regrowth in the distribution system, rather than to benefit water treatment, as the groundwater source is high quality and partially protected (a mix of deep and shallow aquifers). Managers reported that when low-chlorine alarms sound, staff would follow documented procedures to detect any problems, and production would be stopped if the problem could not be corrected. Managers stated that since most processes are automated and monitored at a central control station, it would normally take approximately half an hour to visually check the problem.

To confirm the self-reported benefits and process improvements, quantitative data that potentially captured these changes were analyzed. Since the primary WSP critical control is for free chlorine, the frequency and duration of low-chlorine events at a single production facility were examined using both water quality data sets and alarm logs. The free chlorine readings at the plant outlet were considered out of compliance when below 0.05 mg/l. The data cleaning procedure identified 12 events tied to verified plant shutdowns or sensor failures for removal, resulting in 3% missing data. After cleaning, the number of individual sensor readings was 295,664 in the before period (before 13 November 2012) and 418,027 in the

Table 2		
Readings out of compliance, low-chlorine events, average free chlorine concentration, and event duration before and after WSP implementation (2010–2017), after data cleaning and not considering the 'during' time period from team formation to certification		
Parameter	**Time period**	
	Before	**After**
Total readings	295,664	418,027
Non-compliant readings	10,153	1300
% Non-compliance	3.43%	0.31%
Non-compliance events	1006	179
Average concentration (mg/l)	0.023	0.026
T test for average concentration	1.99 ($p = .047$) (df = 1157)	
Minimum duration (min)	5	5
Maximum duration (h)	22.3	12.1
Median duration (min)	50	5
Average duration (min)	50.6	36.4
T test for average duration	−2.30 ($p = .022$) (df = 1183)	
WSP: Water Safety Plan.		

after period (since 20 December 2013). Statistical analysis comparing the two time periods demonstrated a reduction in the number of low-chlorine events, maximum duration, and average duration, from 1006 events lasting 51 min on average before the WSP to 179 events lasting 36.4 min on average after WSP implementation ($p = .022$; Table 2). Chlorine concentrations during the low-chlorine events showed a small but statistically significant increase from 0.023 to 0.026 mg/l ($p = .047$; Table 2).

After implementation of the WSP, staff members began tracking several additional performance indicators across the full service area (Table 1), which is a common effect of WSP programs.[12] Beyond the chlorine control, other goals of the WSP included monitoring drinking water quality at production, in storage tanks, and in the distribution network around the clock (Table 1). While all performance measures were considered for analysis, many showed discontinuity or little variation.

Those that were continuously tracked were assessed graphically and/or statistically. Indicators showing the most change over time included completion of public utility declarations (objective 1.1), production alarms for low chlorine (objective 2.3), distribution alarms for low chlorine (objective 3.3), rate of sensor availability (objective 3.4), and customer complaints for the full service area (objective 3.5) as well as the single facility's service area (Table 1).

The number of chlorine-related alarms in the production (objective 2.3) and distribution systems (objective 3.3) was tracked, and alarm logs were analyzed for long-term trends. For production, this data tracking method (including 264 alarms across all facilities) was less inclusive and showed fewer events than the high-resolution water quality data from one facility (Table 2). Including all production facilities, Poisson regression with a natural log link function controlling for seasonality showed a significant

decrease in the number of alarms over time ($\beta = -.026$; $p < .001$), with the data leveling out at fewer than five alarms per month (Figure 1). In both production and distribution systems, very few of the alarms that occurred were ultimately verified, since most were due to sensor malfunction. Too few distribution system alarms occurred to enable modeling. Still, records suggested these alarms facilitated beneficial checks and corrections when needed.

Customer complaints for water quality reasons (objective 3.5) revealed an improvement in recording from the onset of WSP implementation in 2013. Prior values from 2012 or earlier were not recorded consistently and thus were excluded from the statistical analysis. Water quality–specific customer complaints in the full service area (e.g. reddish water, chlorine taste, and turbidity) decreased significantly over time ($\beta = -.0071$, $p < .0001$), a change of about 10% fewer complaints per year on

Figure 1

Record of low-chlorine alarms at all production facilities (objective 2.3) from the onset of the WSP implementation period, including a linear trend line
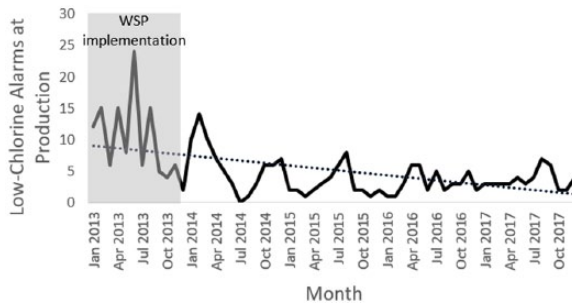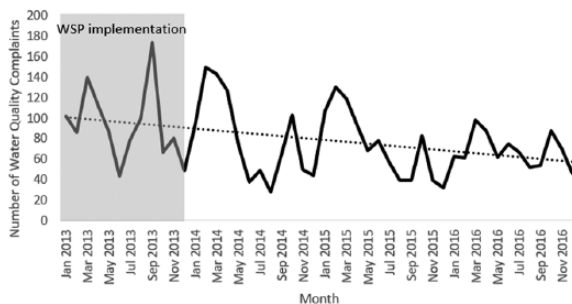

Figure 2

Monthly customer complaints for water quality reasons recorded in the full service area (objective 3.5) from the onset of WSP implementation in 2013; The values of 2012 and before were not consistently recorded

DISCUSSION

Operational improvements stemming from WSPs and reported in the literature have included changes to organizational structure or daily procedures, better risk awareness among water operators, improved water management efficiency, and improved compliance with regulations.[8] Common benefits reported by Suez utility managers across 20 locations included better hazard control (especially for hazards that were previously overlooked), improved treatment practices, and better record keeping.[13] Some of these reported benefits were confirmed by later research, for example, significantly improved compliance with regulatory and internal water quality limits at four of five locations examined.[11] In agreement with the literature, this study focusing on one location demonstrated a reduction in the number of production alarms, shorter non-compliance event duration, fewer customer complaints about water quality, and an improvement in the rate of sensor availability tied to the critical controls prioritized by the WSP. Operational improvements are likely some of the first observable outcomes of WSPs,[3,12] and tracking these early successes may be helpful in reinforcing continual effort toward the WSP.

Improved customer satisfaction may or may not affect revenues but is vital to a utility's reputation and community standing, especially for private companies.[17,18] Improved client and health agency confidence was previously reported by utility managers at 70% of Suez's WSP implementation sites,[13] whereas the site examined in this study noted changes in handling of complaints but did not necessarily perceive greater public confidence. Triangulation with other data types was useful in refining and validating self-reported qualitative outcomes. For example, quantitative analysis in this single-site study documented a statistically significant multiyear decrease in customer complaints for water quality reasons (Figure 2), supporting a potential improvement from the consumers' perspective. It might be valuable to ensure through validation techniques (e.g. third-party customer surveys) that customer or

average (Figure 2). The data suggested complaints typically peaked early in the year with a smaller spike in autumn, and within-year variability appeared to narrow over time (Figure 2). Water quality–specific complaints were also assessed for the single facility alone. Peaks in complaints were visible during the early part of each year (late winter to early spring) (Figure 3). When modeled using Poisson regression with a natural log link function, and controlling for seasonality, complaints decreased significantly over time ($\beta = -.0089$, $p = .0008$). Most of the decline took place in the first year, during the initial WSP implementation period (2013).

Recorded rates of real-time water quality monitoring sensor availability (objective 3.4) reported by the utility (Figure 4) show inter-annual differences, suggesting a marked improvement upon WSP implementation. The apparent slippage in 2016 was likely due to a change in the assessment method to enable improved interpretation, after which the rate improved again in 2017. Prior to 2016, raw monthly data counting the number of continuously working sensors were used, and subsequently the daily data were processed to remove inaccuracies due to periods of maintenance or expected offline activities, to more accurately measure availability. Reported rates of public utility declaration completion (objective 1.1) were fairly consistent over time (Figure 5), suggesting stability or subtle improvement since 2010. This formal documentation serves to confirm the public benefit of a project under French law and is externally tracked but relevant to internal WSP goals regarding documentation and pollution control (Table 1).
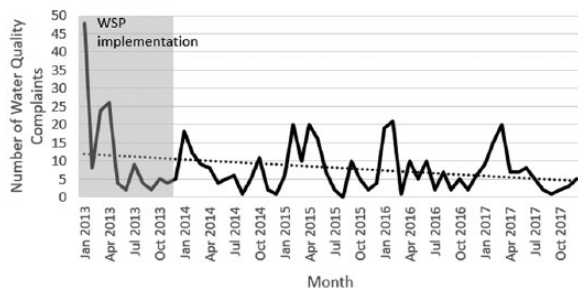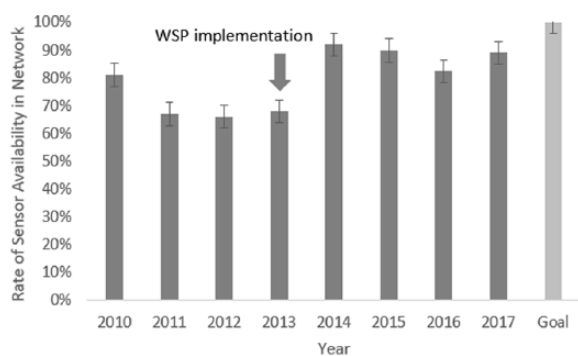
Figure 3

Monthly water quality–specific customer complaints for a single facility recorded from the onset of WSP implementation in 2013, including linear trend line

Figure 4

Comparison of the rate of real-time water quality monitoring sensor availability (objective 3.4) reported each year before, during, and after WSP implementation

employee complaints do not decline due to perceived inability to communicate with the utility or achieve change.[4]

WSPs can potentially permeate all aspects of organizational behavior, including a culture shift toward an improved customer service mentality.[19] Omar et al.[20] showed that incentives used as part of a WSP in Uganda directly linked a renewed management emphasis on customer service to staff rewards, including delegation of responsibilities, financial bonuses, and recognition for meeting goals. Gunnarsdóttir et al.[21] noted that recognition of good staff performance may be a 'bonus' outcome of WSP implementation. When staff felt more in control of the situation, they were less concerned that something might go wrong, making their jobs less stressful. Reduced stress among utility managers

after WSP implementation was similarly noted in the five-country study by Loret et al.[13] However, an improved performance culture may not be a ubiquitous outcome across settings, as WSPs in India and Jamaica suffered from staff perceptions that high standards were unrealistic for low-income countries.[20] The area-wide reduction in customer complaints demonstrated in our study may have stemmed in part from the onset of data recording, which could have increased visibility of performance tracking among staff.

While customer satisfaction represents a key indicator of the proactive organizational culture desired under WSPs, customer concerns (e.g. taste, odor, color) may not directly reflect public health protection. Furthermore, they may sometimes mislead consumers to use

alternative sources that are less safe.[20] For this reason, public health outcomes should be evaluated in tandem with performance indicators such as customer satisfaction.[11] This is especially important where multiple stakeholder goals may exist (e.g. reducing chlorine taste and odor versus maintaining control of potential microbial contaminants).[22] The operational performance and water quality compliance improvements observed in this study did not correspond to a reduced risk of acute gastroenteritis in the service area,[11] although the changes may have altered risks posed by turbidity levels.[6] Cyclical monitoring and evaluation in concert with external experts such as public health agencies can help ensure that well-intended efforts do not have unintended consequences,[23] and assist iterative improvement of the WSP.
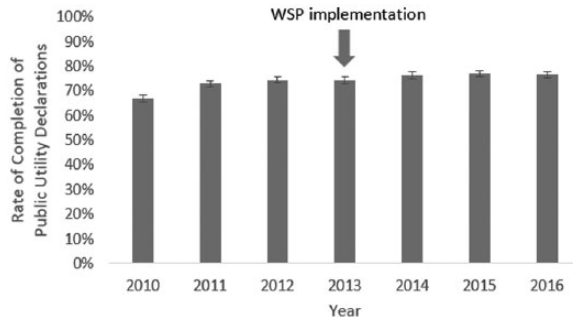
Lockhart et al.[4] provided a taxonomy of WSP performance indicators across outcome categories and recommended 'implementation of improved procedures' as one evaluation category under operational outcomes for WSPs. This includes changes in customer complaints over time and frequency of key operations, such as inspections. Collecting documentation of streamlined procedures was also recommended,[4] although this may or may not indicate how well or how frequently the planned procedures were later carried out. We overcame this challenge by matching self-reported qualitative performance improvements with quantitative water quality and performance records. The improved reaction time reported by utility managers implementing WSPs led to quantification of time needed to return to a state of compliance with critical controls,[13] a central component of WSPs and other risk management programs.

Limitations and recommendations
This study collated operational performance data sources that were not previously assessed independently or reported. Most of the performance indicators were collected beginning in 2012, so earlier records are lacking to make longer-term comparisons. Background rates of change for a larger

Figure 5

Comparison of the rate of completion of public utility declarations (objective 1.1) each year before, during, and after WSP implementation (2017 data not yet available)

indicators grows, a recommended starting point (e.g. list of indicators) for tracking gains would help enable scale-up of progress analyses to the company-wide, national, and/or international level.[2,4,5,13,26] Indicators should consider the resource level of the utility and be integrated into the WSP as monitoring and evaluation targets from the onset of development.[25] Quantitative measures should be supplemented by qualitative 'ground-truthing' involving local utility staff, as direction change could be indicative of underlying improvements in data tracking or changes to external policies.[12]

service area or a different component of the water system (where no WSP was active) would be valuable for isolating causal relationships among observed changes. Some indicators were initiated or dropped during the study period (Table 1), making it difficult to assess trends. While the indicators assessed here were consistently quantified and varied enough to enable some degree of trend analysis, they were largely self-reported historical data that could be affected by reporting bias or social desirability bias. Other means of data collection might include in-depth ethnographic observation or employee or customer surveys, independent of the management team. The customer service culture, for example, could be independently validated through third-party qualitative analysis of recorded calls or call logs.

Furthermore, records from automated water quality sensors likely included some false readings due to sensor drift, failure, or continued recording during plant shutdowns. This posed a challenge to the statistical analysis in determining which types of events to censor, as no strict rule could be definitively applied to all cases. Since automated monitoring approaches may become the standard in the future, it would benefit prospective studies for such data logs to be reviewed and flagged within days of recording. In this study, most data were tracked independently by the utility on a daily or monthly basis and aggregated annually to enable multiyear comparison using an Excel spreadsheet. The WHO offers an Excel quality assurance

tool to document steps taken to complete WSPs, which could potentially be upgraded to incorporate recommended or site-specific evaluation criteria, as well as comparison with historical data from the same location or aggregated data from other locations. More sophisticated software programs might require site-specific development for a given WSP context and would likely need to be custom-built at the utility's expense.

This study demonstrated the need to apply multiple performance measures in combination, since multiple goals may need to be optimized to achieve satisfactory use of the WSP. For instance, achievement of a high degree of personnel involvement may need to be balanced with financial goals seeking to limit expenditures.[13] Utility managers may also be expected to optimize multiple public health goals at the same time, such as short-term acute gastroenteritis risk and long-term bladder cancer risk.[11,16] Establishing some consistency in WSP indicators across implementation sites would help to enable larger scale comparisons; however, some site-specific evaluation criteria may still be warranted. Just as complex public health interventions have core components as well as an 'adaptable periphery',[24] some evaluation criteria may be relevant to only a subset of implementation sites depending on the water source, water system, regulatory environment, and management approach. As knowledge of WSP performance and performance

CONCLUSION

This study validated WSP benefits reported in the literature and demonstrated some potential indicators for consistent application to a broader array of implementation sites, potentially within or across utility groups at a regional-to-global scale. These included the number of critical control alarms (including false alarms), alarm response time (time spent out of compliance with regulatory or internal performance thresholds), the rate of manual monitoring data or water quality sensor availability, the number of water quality–related customer complaints, and progress toward documentation. Initial improvement may be followed in some cases by slippage, and regular monitoring and evaluation cycles can help to ensure iterative improvement of the WSP. Developing greater consistency in the WSP evaluation indicators used across implementation sites can help to clarify synergies and facilitate larger scale progress evaluations.

ETHICAL APPROVAL

In accordance with Committee on Publication Ethics (COPE) requirements, we acknowledge that human subjects research was approved by the Institutional Review

CONFLICT OF INTEREST

FUNDING

SUPPLEMENTAL MATERIALS

Requests for raw data or statistical code should be submitted to the corresponding author.

ORCID ID

Karen Setty https://orcid.org/0000-0002-5591-9693

References

1. World Health Organization. *Guidelines for drinking-water quality*. 3rd ed. vol. 1. Geneva: WHO; 2008.
2. WHO and IWA. *Global status report on Water Safety Plans: a review of proactive risk assessment and risk management practices to ensure the safety of drinking-water*. Geneva: WHO; 2017.
3. Gelting RJ, Delea K, Medlin E. A conceptual framework to evaluate the outcomes and impacts of Water Safety Plans. *J Water Sanit Hyg Dev* 2012;**2**:103.
4. Lockhart G, Oswald WE, Hubbard B *et al*. Development of indicators for measuring outcomes of Water Safety Plans. *J Water Sanit Hyg Dev* 2015;**4**:171–81.
5. Vieira JMP. A strategic approach for Water Safety Plans implementation in Portugal. *J Water Health* 2011;**9**:107–16.
6. Setty KE, Enault J, Loret J-F *et al*. Time series study of weather, water quality, and acute gastroenteritis at Water Safety Plan implementation sites in France and Spain. *Int J Hyg Environ Health* 2018;**221**:714–26.
7. Jetoo S, Grover VI, Krantzberg G. The Toledo drinking water advisory: suggested application of the water safety planning approach. *Sustainability* 2015;**7**:9787–808.
8. Kot M, Castleden H, Gagnon GA. The human dimension of Water Safety Plans: a critical review of literature and information gaps. *Environ Rev* 2015;**23**:24–9.
9. String G, Lantagne D. A systematic review of outcomes and lessons learned from general, rural, and country-specific Water Safety Plan implementations. *Water Sci Technol Water Supply* 2016;**16**:1580–94.
10. Gunnarsdóttir MJ, Gardarsson SM, Elliott M *et al*. Benefits of Water Safety Plans: microbiology, compliance, and public health. *Environ Sci Technol* 2012;**46**: 7782–89.
11. Setty KE, Kayser GL, Bowling M *et al*. Water quality, compliance, and health outcomes among utilities implementing Water Safety Plans in France and Spain. *Int J Hyg Environ Health* 2017;**220**:513–30.
12. Kumpel E, Delaire C, Peletz R *et al*. *Water Safety Plan impact assessment in the Asia Pacific region*. World Health Organization; 2016.
13. Loret JF, Thé CB, de Alonso JM *et al*. *Assessing the costs and benefits of Water Safety Plans. Water Safety Portal (wsportal. org)*, 2016. Available online at: http://www.wsportal.org/resource/assessing-the-costs-and-benefits-of-water-safety-plans/
14. Morse JM. Principles of mixed methods and multimethod research design. In: Tashakkori A. Teddlie C (eds) *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage; 2003. pp. 189–208.
15. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs – principles and practices. *Health Serv Res* 2013;**48**:2134–56.
16. Corso M, Galey C, Beaudeau P. *Évaluation Quantitative De L'impact Sanitaire Des Sous-produits De Chloration Dans L'eau Destinée À La Consommation Humaine En France*. Paris: Santé Publique; 2017.
17. Doria França De M. Factors influencing public perception of drinking water quality. *Water Policy* 2010;**12**:1–19.
18. Hrudey SE, Hrudey EJ, Pollard SJT. Risk management for assuring safe drinking water. *Environ Int* 2006;**32**:948–57.
19. Summerill C, Pollard SJT, Smith JA. The role of organizational culture and leadership in Water Safety Plan implementation for improved risk management. *Sci Total Environ* 2010;**408**:4319–27.
20. Omar YY, Parker A, Smith JA *et al*. Risk management for drinking water safety in low and middle income countries – cultural influences on Water Safety Plan (WSP) implementation in urban water utilities. *Sci Total Environ* 2017;**576**:895–906.
21. Gunnarsdóttir MJ, Gardarsson SM, Bartram J. Icelandic experience with Water Safety Plans. *Water Sci Technol* 2012;**65**:277–88.
22. Summerill C, Smith J, Webster J *et al*. An international review of the challenges associated with securing 'buy-in' for Water Safety Plans within providers of drinking water supplies. *J Water Health* 2010;**8**: 387–98.
23. Bereskie T, Rodriguez MJ, Sadiq R. Drinking water management and governance in Canada: an innovative Plan-Do-Check-Act (PDCA) framework for a safe drinking water supply. *Environ Manage* 2017;**60**:243–62.
24. Proctor E, Silmere H, Raghavan R *et al*. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011;**38**:65–76.
25. Mudaliar MM. Success or failure: demonstrating the effectiveness of a Water Safety Plan. *Water Sci Technol Water Supply* 2012;**12**:109–16.
26. Kumpel E, Delaire C, Peletz R, *et al*. Measuring the Impacts of Water Safety Plans in the Asia-Pacific Region. *Int J Environ Res Public Health* 2018; 15: 1–18.