Waterlines, 41:1, 5–23
<http://dx.doi.org/10.3362/1756-3488.18-00008OA>

Measuring progress towards sanitation and hygiene targets: a critical review of monitoring methodologies and technologies

Nick Turman-Bryant, Thomas F. Clasen, Kathryn Fankhauser, and Evan A. Thomas

Abstract: The Sustainable Development Goal (SDG) target for access to safe sanitation and hygiene represents a marked improvement over the target used during the Millennium Development Goal (MDG) period. The SDG target attempts to: explicitly address hygiene; eliminate inequalities within populations; evaluate sanitation services beyond the household; account for the accessibility, safety, acceptability, and affordability of service delivery; and improve the sustainability of services (WHO/UNICEF, 2015). However, the proposed indicators for monitoring progress in sanitation and hygiene still rely primarily on infrequent household surveys and census data. This paper provides a critical review of the sanitation and hygiene target and explores the potential gaps between the expanded understanding of access, the proposed monitoring strategies, and the desired impacts. A variety of innovative methodologies and technologies are reviewed, with specific attention given to their suitability for measuring and monitoring progress towards the sanitation and hygiene target.

Keywords: Sustainable Development Goals, sanitation, monitoring technologies, indicators, causal modelling

Introduction

Sustainable Development Goal (SDG) Target 6.2 is: 'By 2030, achieve access to adequate and equitable sanitation and hygiene for all and end open defecation, paying special attention to the needs of women and girls and those in vulnerable situations' (WHO/UNICEF, 2015). While the corresponding Millenium Development Goal (MDG) target emphasized a single outcome – access to improved sanitation facilities – the SDG sanitation target builds on this by incorporating adequacy and equity. By including 'for all', the target mandates that sanitation systems and services be available to all people at all times, regardless of age, gender, disability

Nick Turman-Bryant (nicktb@pdx.edu), Portland State University, USA; Thomas F. Clasen (thomas.f.clasen@emory.edu), Emory University, USA; Kathryn Frankhauser (fankhaus@ohsu.edu), Oregon Health and Science University; Evan A. Thomas (evan.thomas@colorado.edu), Mortenson Center in Engineering for Developing Communities, University of Colorado at Boulder, USA

status, or income level. Furthermore, the incorporation of child faeces disposal into the definition of open defecation requires that all faeces be disposed of in a safe and hygienic manner, whether in an improved sanitation facility or treatment system. Lastly, the addition of special attention to women, girls, and those in 'vulnerable populations' requires that additional measures be met to provide for the special sanitation needs of women and girls, as well as to ensure that all people in 'refugee camps, detention centers, mass gatherings, and pilgrimages' have adequate sanitation.

This expanded understanding of access recognizes that access should not be reduced to the binary measure of improved or unimproved facility type. Access can change over time, is influenced by structural and relational mechanisms, and operates on varying scales. All of these facets of access interact dynamically to influence the ability of individuals, households, and communities to derive benefits from sanitation and hygiene services (Ribot and Peluso, 2003). Although this expanded understanding of access represents an improvement over the previous binary definition, a gap persists between the stated desire to improve service levels (e.g. to 'promote progressive improvements in the quality of services based on the normative criteria of the human right to water and sanitation i.e. accessibility, quantity, quality, acceptability, and affordability') and the proposed monitoring strategies (WHO/UNICEF, 2015: 7). For example, household surveys and spot-checks are often the primary indicators for measuring progress, despite being weak proxies for facility use and safety (Clasen, 2018). The combination of indicator uncertainty with an extended chain of inference (i.e. facility type and self-reported use as proxies for safety and actual use, and safety and actual use as proxies for health impact) makes monitoring progress on the SDG target difficult and contributes significantly to the overall uncertainty associated with evaluations of health impact.

After briefly illustrating how measurement uncertainty combines with inferential uncertainty in causal modelling, this paper then provides a critical review of the proposed sanitation and hygiene service ladders in relation to an expanded notion of access and service levels (i.e. accessibility, safety, use, equity, and acceptability). Finally, a variety of innovative methodologies and technologies for monitoring are reviewed, with a discussion of their relative strengths and weaknesses.

Causal modelling and indicator selection

In *Evidence-Based Policy,* Cartwright and Hardie describe how causal models are composed of a constellation of multi-faceted conditions that contribute to a desired effect (Cartwright and Hardie, 2012). While individual conditions are sufficient to cause an impact, they are considered unnecessary because of the variety of conditions that can produce the desired effect (e.g. improved sanitation, water, or hygiene services). However, conditions are themselves composed of parts that are insufficient but non-redundant; that is, each part is unable to cause an impact on its own but it is a necessary component of the overall condition that contributes to the desired effect. Figure 1 provides a visual depiction of these insufficient but necessary

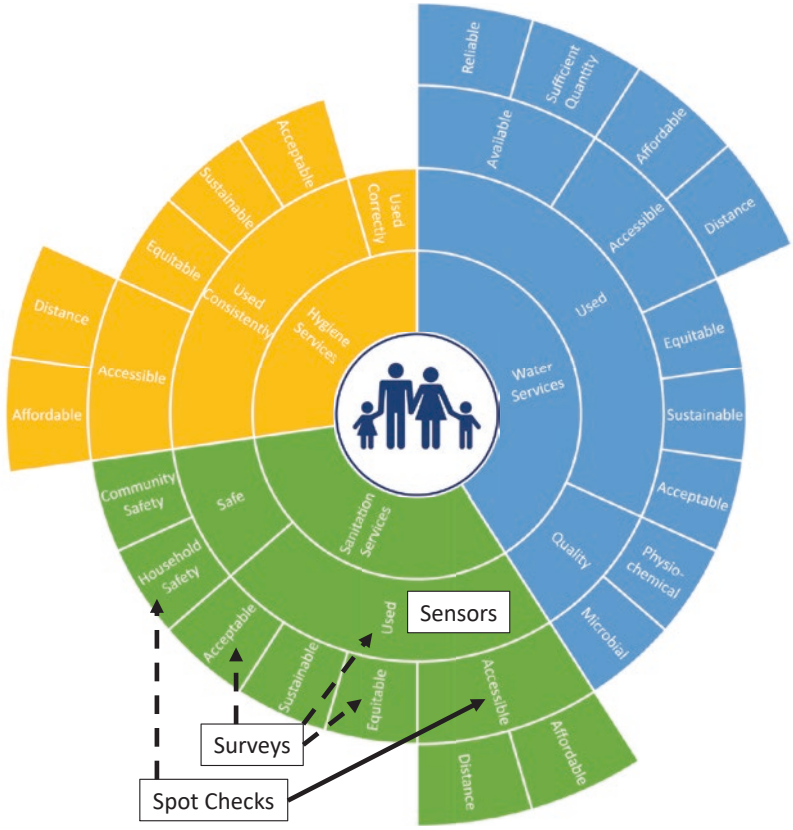

Figure 1 Diagram mapping the conditions and indicators for measuring impact in water, sanitation, and hygiene

components of unnecessary but sufficient (INUS) conditions, where the sufficient conditions for health impact are represented as rings, and each ring segment represents a non-redundant component of that condition. This figure illustrates that all components of a condition must be monitored in order for the condition to be sufficient, and that two of the more common methods used for monitoring – spot-checks and household surveys – serve as weak proxies for actual impact. For example, sanitation services will only have an impact on health if they are safe and if they are used. Similarly, sanitation services will only be used by all members of a household if they are acceptable, accessible, equitable, and sustainable. As a result, a spot-check can serve as a strong indicator of the accessibility of a facility, because it is visually verified, as opposed to household surveys that often rely on self-reporting to determine how often a facility is used, who uses it, and whether it is acceptable. Similarly, facility type may be a weak proxy for the safety of the facility for the household because there is no objective verification of how effectively the facility is separating faecal waste from human contact. Also, without any indicators for use or faecal sludge management, the assumption of health impact is dubious.

However, uncertainty in inference can be reduced by using stronger monitoring indicators that are closer to the desired effect.

This point is echoed in a recent systematic review of indicator selection methods for water, sanitation, and hygiene (WASH) monitoring, where Schwemlein et al. (2016) note that there is a general lack of consistency, specificity, and relevancy in the indicators used by the projects and programmes included in their review. In particular, they suggest that better coordination of WASH indicators could help 'identify weaknesses in data collection', 'inform decisions in WASH policy and practice', and 'facilitate comparison of projects, programs, and interventions' (ibid.: 2). However, Schwemlein et al. argue that a more formal process for selecting indicators and organizing data collection is needed to improve transparency and improve coordination in WASH interventions. Notably, they recommend that the indicator selection process should be explicitly tied to the outcomes of interest, based on the purpose and scope of the intervention. Finally, they suggest that proposed indicators should be evaluated using objective selection criteria, including whether the proposed indicator is measureable, reliable, and sensitive to changes in the outcome of interest. They also argue that candidate indicators must be valid, that is, '[t]here must be an accurate correlation between an indicator and the issue for which it is supposed to proxy', based on existing data (ibid.: 11).

Sanitation and hygiene service ladders

Recognizing that sanitation services can include a variety of levels, the Joint Monitoring Program (JMP) has updated its service ladder to define five thresholds. Like the service ladder used for the MDG sanitation target, three of the categories designate the type of sanitation facility: unimproved, limited, or basic, where 'limited' refers to latrines that are shared by two or more households. 'Open defecation' describes the deposition of human faeces directly in the environment, and 'safely managed' designates a basic sanitation facility that is not shared and where excreta are safely disposed of in situ or treated off site. Similarly, the hygiene service ladder is primarily concerned with the presence of handwashing facilities on the premises, where basic facilities have soap and water and limited facilities have no soap or water. Figure 2 maps these categories from the proposed sanitation and hygiene service ladders to the outcomes identified in the target descriptions (WHO/UNICEF, 2015). Although a variety of outcomes are recognized in the target descriptions (i.e. accessibility, safety, equity, and acceptability) and use is implied, many of these components are not explicitly represented in the proposed service ladders.

ACCESSIBLE: Accessibility is defined as 'facilities that are close to home that can be easily reached and used when needed' (WHO/UNICEF, 2015: 11). Accessibility is well represented in both service ladders, with access being inferred directly from the observation of sanitation and hygiene facilities on the premises.

SAFE: The safety of the sanitation facility for the household – how well it separates excreta from human contact within the household – is not represented in the ladder. Household safety is indirectly inferred from the sanitation facility type,

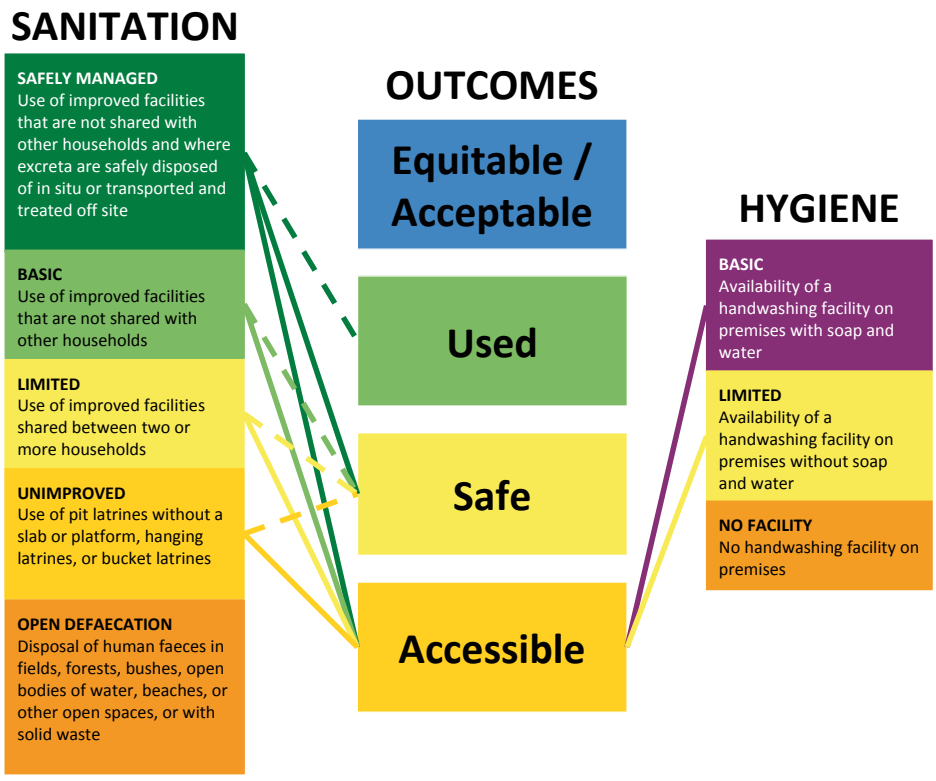

Figure 2 Mapping the categories of the sanitation and hygiene ladders to the desired outcomes. Line type designates the strength of the representation from each category to each outcome

where basic sanitation facilities are assumed to adequately separate excreta from household contact, and unimproved sanitation facilities are not. The safety of the sanitation facility for the community – how well it separates excreta from human contact beyond the household – is better represented in the ladder. Like household safety, community safety is inferred indirectly from the sanitation facility type, but community safety is addressed directly by the 'safely managed' category.

USED: Although sanitation and hygiene facility use is key for realizing health benefits, use is not addressed explicitly in any of the service categories. With respect to hygiene, it is also helpful to distinguish between adherence and technique, where technique designates efficacy in removing contamination, and adherence designates the consistency of use. Although none of the service categories address handwashing technique or adherence, regular and effective handwashing that coincides with sanitation behaviours is also important for realizing health benefits.

EQUITABLE/ACCEPTABLE: Equity is defined in the target as the 'progressive reduction and elimination of inequalities between sub-groups' (WHO/UNICEF, 2015: 11), but neither equity nor acceptability are represented explicitly in the sanitation ladder. While population-level inequalities could be inferred based on

adoption rates or the elimination of inequalities across population sub-groups, more direct measures of equity and acceptability may be needed to capture intra-household use and the acceptability of specific sanitation and hygiene interventions. For example, the needs of women and girls require special consideration, given females' higher risk of experiencing harassment and violence when safe and private sanitation facilities are not available (Bangdiwala et al., 2004; SHARE, 2014; Sclar et al., 2018). Beyond considerations of equity, acceptability could depend on the desirability of the technology and whether it is aspirational. In either case, if reliable measures of use are available, it is possible that equity and acceptability could be inferred indirectly from sanitation and hygiene facility use.

Sanitation and hygiene indicators

Although the inclusion of the 'safely managed' category represents substantial progress in the evolution of the service ladder, there is still a great deal of ambiguity surrounding the indicators that will be used to monitor outcomes. Currently, the JMP plans on using household surveys and regulatory data as the main data sources for observing household sanitation facility types. Sanitation facility type and attributes will then be used to infer other outcomes like safety and use. For example, lack of use could be inferred directly from a non-functioning toilet, and unimproved sanitation facilities would be assumed to provide unsafe management both within and beyond the household (WHO/UNICEF, 2015).

Given that hygiene was not addressed in the MDG targets, its inclusion in SDG 6.2 highlights the growing consensus that water, sanitation, and hygiene are indelibly linked and cannot be treated in isolation. Hygiene is defined as 'the conditions and practices that help maintain health and prevent disease including handwashing, menstrual hygiene management, and food hygiene' (WHO/UNICEF, 2015: 11). While food hygiene was identified as one of the top priorities for health and nutrition, it was ultimately determined to be outside the scope of WASH monitoring. Similarly, menstrual hygiene management is mentioned but is not addressed specifically in the indicators. Thus, hygiene as described by the service ladder is primarily concerned with handwashing. Like the sanitation ladder, the hygiene service ladder uses the presence of a handwashing facility as a proxy for quality and use.

Sanitation and hygiene beyond the household

One clear advance of the proposed indicators is the focus on faecal sludge management. The indicator defines 'safely managed sanitation' as systems in which faecal waste is transported through a sewer to a designated location, is collected from systems by a process that limits human contact, and is transported to a designated location or undergoes, as a minimum, secondary treatment or 'primary treatment with long ocean outfall for sewerage' or is treated at a 'managed disposal site' or waste water treatment plant or 'stored on site until … safe to handle and re-use' (WHO/ UNICEF, 2015: 28). This indicator is designed to encompass essential services and operational requirements for public health benefits (Feachem et al.,

1983; Shuval, 2003; Escamilla et al., 2013). At the same time, the indicator does not evaluate the integrity of the system or services; neither is there consideration of sustainability, where sustainability is defined as the ability to sustain 'services to ensure lasting benefits' by 'safely managing human waste' (WHO/UNICEF, 2015: 8). Also, given the higher infectivity of child faeces, special consideration should be given to their disposal and management (Rand et al., 2015; WSP/UNICEF, 2015).

The JMP indicators also acknowledge that monitoring the safe management of excreta requires a full faecal waste flow framework that spans the service chain from containment to reuse or disposal. While information about containment can be collected from household surveys, the JMP proposes to monitor the emptying, transport, and treatment of faecal waste using a combination of utility, population, and household data to estimate safe management through the service chain. As a result, 'reuse and disposal would not be monitored initially at a global level' (WHO/ UNICEF, 2015: 28). On-site treatment and disposal would be inferred based on a variety of factors, including the sanitation facility type, construction quality, frequency of use, population density, geographic conditions, and urban versus rural location (WHO/UNICEF, 2015). Off-site treatment will initially be estimated from utility records, based on the number of sewer connections and installed treatment facilities. Off-site treatment for excreta that are collected and transported from septic tanks and pit latrines could then be estimated using records from trucks disposing waste at waste water treatment plants (WHO/UNICEF, 2015).

Monitoring sanitation and hygiene beyond the household also entails the evaluation of institutions such as schools and health clinics, where facility use and the risk of exposure to faecal pathogens are high. The negative impacts of incomplete sanitation coverage at the community level have been documented in field studies and systematic reviews (Moraes et al., 2004; Barreto et al., 2007; Geruso and Spears, 2015). A study of city-wide sanitation improvements in Salvador, Brazil, saw overall reductions in the prevalence of diarrhoea by 21 per cent, and in high-risk areas with high baseline prevalence the reduction was 43 per cent (Barreto et al., 2007). The new guidelines define three simplified service ladders (basic, unimproved, and no service) for sanitation, handwashing, and menstrual hygiene facilities. Institution-based indicators will rely on a combination of sector-based management information systems (MIS), site surveys of facilities, and household surveys that incorporate questions about extra-household facility access and use (WHO/UNICEF 2015). In addition, the JMP acknowledges the need to monitor WASH services for vulnerable populations (e.g. refugee camps or detention centres), and data will primarily be sourced from surveys conducted by the United Nations High Commissioner for Refugees (UNHCR) and the International Committee for the Red Cross (ICRC) for global reporting (WHO/UNICEF, 2015).

Monitoring sanitation outcomes

The following section provides an overview of relevant practices and technologies for monitoring sanitation outcomes. As no one practice or technology is adequate for monitoring progress in sanitation, it is important to note that some

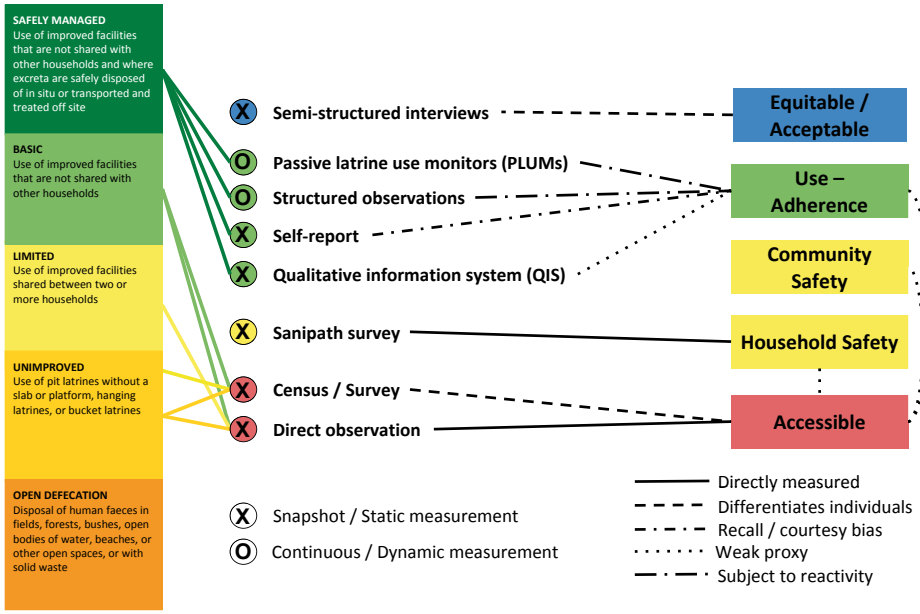

Figure 3 Monitoring methodologies and technologies for sanitation outcomes

practices and technologies are better suited for monitoring specific sanitation outcomes. Figure 3 provides a visual mapping of each methodology to each sanitation outcome.

Accessible

Household surveys and national censuses are the most common methodologies used for assessing a household's access to sanitation facilities (Clasen et al., 2012). There are a variety of advantages to using surveys for evaluating access. First, as one of the most common tools for gathering household information, there is a growing knowledge base that facilitates comparison across time and geography. Second, appropriate survey design can result in higher validity and reliability of survey responses. Third, administering surveys in households allows for interaction with household members and direct observation of sanitation facilities. Thus, while survey questions can differentiate which members of the household are able to access a sanitation facility, direct observation allows an observer to verify the sanitation facility type, its functionality, whether it is private or shared, and its proximity to the household. Still, unless there are repeated visits to the household, census surveys and spot-checks only provide a static measurement of the sanitation facility's accessibility and functionality (Thomas and Mattson, 2013).

Household safety

The type of sanitation facility and whether it is private are the two main proxies used to determine whether excreta are safely separated from human contact

within the household. However, there are very few methods for directly measuring the quality of specific sanitation facilities. One exception is the Sanipath Rapid Assessment Tool created by researchers at the Center for Global Safe Water at Emory University in the USA. The Sanipath tool provides an assessment of exposure to faecal contamination by measuring the level of faecal contamination associated with different transmission pathways (e.g. drinking water, latrines, produce, open drains, etc.) These microbial loads are combined with surveys that characterize household behaviours to generate risk assessments for each exposure pathway. For example, a household may use a private pit latrine with a slab, but the Sanipath tool could be used to estimate the actual risk of exposure to faecal contaminates based on the level of contamination in the latrine and the behaviours of the users. While the Sanipath tool is primarily designed to evaluate the level of exposure to faecal contamination for an entire community, the methodology could be adapted to the household scale. The ability to combine microbial testing with survey responses is also a strength, as the surveys facilitate a more nuanced characterization of individual sanitation and hygiene practices. However, the tool depends on the ability of local laboratories to conduct testing in a sterile environment with sufficient equipment. Also, unless the Sanipath assessment is performed regularly, the measurement represents a snapshot in time that is not able to monitor changes in behaviour, faecal contamination in the environment, or the functionality of sanitation facilities (Sanipath, 2014).

Community safety

Similarly to the situation for monitoring household safety, there are very few methodologies that have been developed to directly verify the safe management of excreta beyond the household. While safe management is often assumed for sanitation facilities that are connected to a sewer system, septic tank, or pit latrine, the actual verification of waste removal, transport, and treatment represents a significant challenge for monitoring community safety. Data from utilities could be used to estimate the safe treatment of excreta based on the number of household connections and the conveyance to installed treatment facilities. Similarly, records from disposal trucks could be used to estimate the number of households where waste is safely collected and removed (WHO/UNICEF, 2015). However, unless records from the point of collection to the point of treatment can be corroborated, utility and waste removal estimates may underestimate leakage or the deposition of waste directly into the environment.

Although they may not provide an accurate measure of the level of exposure to faecal contamination in the environment, records from utilities and waste collectors can be used to verify that excreta are being collected and conveyed to treatment facilities. For example, Sanergy Inc. in Kenya has partnered with SweetSense Inc. in Portland, Oregon to use motion sensors to optimize sanitation waste collection operations. The sensors are also able to send alerts from the latrine operator or the waste collector through Radio Frequency Identification (RFID) tags that are directly integrated into Salesforce, a logistics and customer management platform.

Similarly, x-runner in Peru and the Water and Sanitation Program in India are able to track the installation and management of improved toilets through Near Field Communication (NFC) tags and Quick Response (QR) codes that are scanned and tracked through Salesforce and Open Data Kit, a mobile survey application (Robiarto et al., 2014; Nique and Smertnik, 2015).

Use – adherence

Although not explicitly represented in the sanitation service ladder, sanitation facility use is a key indicator for measuring sanitation facility efficacy (Clasen et al., 2014). However, the verification of sanitation facility use and a household's adherence to use is incredibly challenging.

BRAC in Bangladesh has developed a Qualitative Information System (QIS) that incorporates a combination of spot check indicators with survey questions to assess latrine use. In a study comparing three latrine utilization methodologies, including surveys, observations, and motion detector sensors, there was a strong correlation between latrine spot check indicators and BRAC's QIS indicators. There was also a positive correlation between self-reported latrine use and sensor-recorded latrine use, although self-reported use was significantly greater than sensor-recorded use. While households reported an average of 32.8 latrine uses over four days, sensors recorded an average of 21.7 uses, perhaps indicating recall or courtesy bias in self-reporting (Delea et al., 2017). Given the different scales used, no comparison was drawn between sensor-recorded use or self-reported use and the spot check indicators.

Clasen et al. suggest that spot check indicators and sanitary surveys 'are subjective and may lack necessary sensitivity and specificity to quantify patterns of use' (Clasen et al., 2012: 3296). In an experiment comparing motion-detector-sensor-equipped latrine use against structured observations, they found that sensor-recorded use and observed use agreed within two latrine use events 93.9 per cent of the time over 228 observation periods. They also found strong evidence of reactivity to structured observation, as the sensors recorded significantly more latrine events during observation periods compared with non-observation periods (Clasen et al., 2012). O'Reilly et al. also recorded a high level of agreement between sensor-reported events and structured observations (O'Reilly et al., 2015).

In a similar study, Sinha et al. found that mean reported 'usual' daily use was almost twice the average daily sensor-recorded use (7.09 versus 3.62 events). While there was better agreement between reported use and sensor-recorded use from the previous 48 hours (4.61 versus 3.59 events), the predicted number of latrine events using the 48-hour recall measure was still 60 per cent greater than the average number of events recorded by the sensors (Sinha et al., 2016).

In this regard, sensors like the passive latrine use monitor (PLUM), provide perhaps the best estimate of actual latrine utilization. They are subject to less reactivity compared with structured observations, can provide higher resolution data over longer observation periods in near time, and can be incorporated unobtrusively in a variety of sanitation facility settings. However, unlike structured observations and self-reporting, PLUMs are not able to differentiate which individuals are using the sanitation facility in the household. Also, while PLUMs have a high degree of

accuracy for household latrines, use of public latrines is more difficult to characterize due to the shorter intervals between latrine uses. As a result, sensors like PLUMs should ideally be combined with surveys or observations to better characterize sanitation facility use and adherence.

Equitable/acceptable

Although elimination of inequalities and the special needs of women and girls are addressed in the sanitation target, the JMP's current proposal for measuring inequalities involves a comparison of population sub-groups that are disaggregated by 'income, gender, age, race, ethnicity, migratory status, disability, geographic location, and other characteristics relevant in national contexts' (WHO/UNICEF, 2015: 17). However, evidence shows that inequalities in sanitation facility use can occur at an intra-household scale as well as a societal scale (Jenkins and Curtis, 2005; Coffey et al., 2014). In addition, more nuanced methodologies may be needed to incorporate the specific needs of women and girls to ensure the acceptability, security, and privacy of sanitation facilities. Given the sensitivity of sanitation subjects and the influences of cultural and religious norms, qualitative methodologies like ethnography and semi-structured interviews may be needed to accurately gauge acceptability and characterize intra-household sanitation behaviours. For example, O'Reilly et al. found that ethnographic and motion detector data were highly complementary and useful for comparing sanitation practices between groups that differed in geography and religious affinity (O'Reilly et al., 2015).

Monitoring hygiene outcomes

In contrast to the challenges associated with monitoring progress on sanitation outcomes, the proposed methodology for monitoring progress on hygiene is relatively direct. Since 2009, the JMP has used the 'observation of the place where household members wash their hands and the presence of water and soap' as the primary indicator of handwashing behaviour (WHO/UNICEF, 2015: 21). As a result, the JMP is able to measure the hygiene service ladder directly through household surveys and extrapolate those estimates to the broader population base.

While the monitoring of hygiene facilities is relatively straightforward, the following section provides a summary of different practices and technologies that have been used to monitor specific hygiene outcomes. Actual handwashing behaviour is still challenging to monitor, but it is possible that the type of hygiene facility could serve as an adequate proxy for access and use for mixed-purpose large-population surveys (Ram, 2013). Figure 4 provides a visual mapping of each methodology to each hygiene outcome.

Accessible

As when evaluating sanitation, household surveys and censuses are the easiest indicators for evaluating access to hygiene facilities. Easily combined with spot check indicators that facilitate direct observation of handwashing facilities and

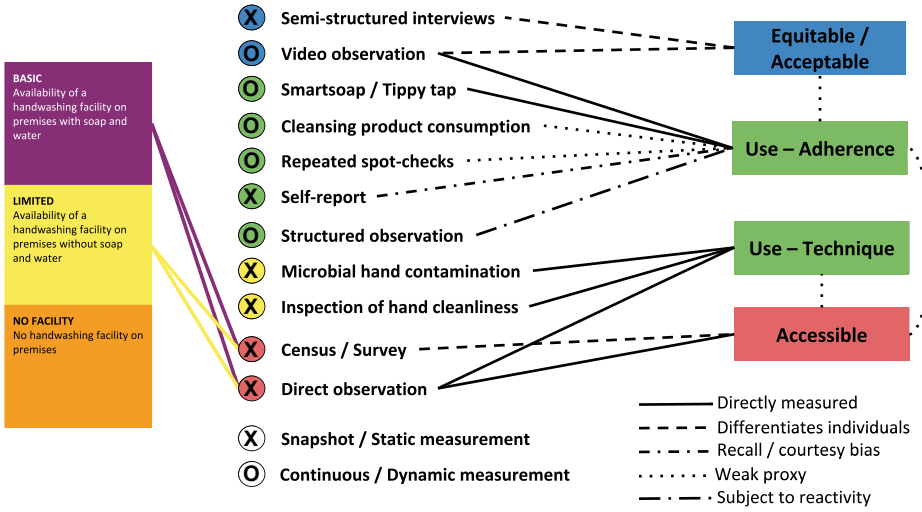

Figure 4 Monitoring methodologies and technologies for hygiene outcomes

materials, rapid observations are used almost exclusively in large-population surveys where hygiene is one among many behaviours of interest. As a direct measure, rapid observations are cost-effective, efficient, and more reliable than survey responses (Cairncross et al., 2005; Ram, 2013). However, verification of the handwashing facility does not provide information about individual hygiene practices within the household, whether handwashing is performed at critical times (e.g. after defecation or before meals), or the efficacy of handwashing and its consistency over time.

Use – technique

Measurement of microbial hand contamination through laboratory measurements or visual inspection are two methodologies that are used to verify handwashing efficacy. While research has shown a positive correlation between hand contamination and health outcomes (Luby et al., 2009; Pickering et al., 2010), measurement of hand contamination is relatively expensive, time-consuming, and may require access to a microbial laboratory facility (Ram, 2013). Observation of handwashing practice can be a useful method for verifying the use of soap, the duration of handwashing, and the method for drying, but respondent behaviour may be influenced by the presence of an observer (Sagerman et al., 2011). Visual inspections of hand contamination can be performed efficiently and are positively associated with microbial contamination and observed handwashing (Pickering et al., 2010). However, a high inter-observer reliability is important for avoiding subjectivity bias between multiple enumerators (Ram, 2013). Also, as a static indicator, the measurement of hand contamination is not able to capture how quickly recontamination occurs after washing. For example, Ram et al. found a high level of recontamination within two hours of a thorough handwashing with soap (Ram et al., 2011).

Use – adherence

Although handwashing with soap at critical times (e.g. after defecation and before meals) has been identified as one of the most cost-effective behaviours for preventing infection, verification of handwashing adherence remains a challenge. Indicators like the presence of soap and water and handwashing efficacy are positively correlated, but it is still unclear how well these indicators predict handwashing behaviour (Ram, 2013). For example, Biran et al. (2008) found that only 2 out of 26 handwashing indicators used to classify households as 'handwashing' – the presence of soap beside the latrine and soap in the yard – were significantly correlated with classifications of households based on structured observations.

Self-reported behaviour is one of the most common indicators used to assess hygienic practice. However, self-reported handwashing usually overestimates actual handwashing due to the social desirability associated with handwashing. For example, while 77 per cent of respondents in a Bangladesh study reported handwashing with soap after defecation, only 32 per cent were observed to do so (icddr,b, 2008). When accounting for actual soap use, the discrepancy between reported and observed handwashing persists but decreases slightly (Ram, 2013).

Structured observation has typically been used as the gold standard for comparing different handwashing measures. Structured observations can record more detailed information about how hands are washed, when hands are washed, and who washes their hands, but it is important that the timing and location of observations include as many members of the household as possible and critical events like meal preparation and consumption (Biran et al., 2008). Given that an observer's presence has been shown to increase the number of handwashing events by as much as 35 per cent, Ram et al. question whether structured observations should be the standard for comparison, due to high reactivity (Ram et al., 2010). Unlike self-reports, however, structured observation provides a more dynamic measure of handwashing behaviour over time. Repeated spot-checks also provide a more dynamic measure of handwashing behaviour. Webb et al. determined that six separate spot-checks are needed to reliably estimate a household's hygiene practices (Webb et al., 2006), although repeated visits may also increase reactivity (Arnold et al., 2015).

While studies that monitor the consumption of cleansing products have been conducted in high-income countries, there are only a few examples of studies that have tracked soap purchases or soap weight differences as a proxy for handwashing behaviour (Ram, 2013). For example, Gadgil et al. found a positive correlation between consumption of bar soap and observed handwashing events (Gadgil et al., 2011). However, Luby et al. observed no differences in soap purchases between the treatment and control groups in a handwashing intervention, despite differences between the two groups in the presence of soap and water and handwashing techniques (Luby et al., 2009).

Sensors can also provide an objective and unobtrusive characterization of handwashing behaviour. For example, SmartSoap is an ordinary-looking bar of soap with an embedded accelerometer that measures motion on three axes. On its own, SmartSoap can provide an accurate count of the number of times the soap bar is used each day, although it cannot differentiate users or determine whether handwashing

coincides with critical events like defecation or meals. However, by combining SmartSoap data with data from a motion sensor placed on the vessel holding water for anal cleansing, Biran et al. were able to detect handwashing events after defecation. Although overall soap use increased, they found that there was no increase in the number of soap uses following defecation (Biran et al., 2008). Similarly, Mercy Corps partnered with SweetSense to use motion sensors with water flow sensors to monitor the prevalence of handwashing after latrine use. They found that water use after latrine use was very low (<10 per cent) in all but one district, which registered almost 40 per cent use of water after latrine use. They also found that self-reported use of the latrine and handwashing after using the latrine were much greater (up to 4 times and 25 times, respectively) than the latrine use and handwashing after latrine use detected by the sensors (Thomas and Mattson, 2013).

Finally, video observation can be an effective tool for observing and recording handwashing behaviour unobtrusively. Although no comparisons have been conducted, it is possible that video observation would be preferable to direct observation in settings where handwashing behaviour can be clearly recorded from a fixed location. Video observation has the advantage of being able to record over longer time periods without interruption, and recordings can be reviewed rapidly by a human observer. It is also possible that discreetly placed video observation may reduce reactivity, although there are ethical concerns that must be considered when consent cannot be obtained for all involved parties. Like sensors, video observation provides a dynamic measure of handwashing behaviour over time, but it also allows the reviewer to differentiate the handwashing behaviours of specific individuals (Pickering et al., 2014).

Equitable/acceptable

Like sanitation, the goals of equity and acceptability in hygiene practices may require more qualitative methodologies like ethnography and semi-structured interviews to understand what motivates hygienic behaviour, to gauge the acceptability of hygiene interventions, and to characterize intra-household hygiene behaviours. This is particularly true for the special needs of women and girls and the ambiguity surrounding indicators for menstrual hygiene management. For example, Curtis et al. found that social affiliation and disgust were two strong motivators of handwashing behaviour, but that fear of disease had little influence on handwashing behaviour (Curtis et al., 2009).

As proposed by Ram, composite measures would ideally be employed to more accurately characterize handwashing behaviours (Ram, 2013). While some methodologies are particularly suited to measuring specific outcomes (e.g. sensors for monitoring handwashing practices), no one methodology is adequate for verifying and monitoring all four hygiene outcomes.

Combined methodologies

Given that all monitoring and evaluation methods have their own advantages and limitations, it is often beneficial to use more than one method to get a fuller picture of WASH behaviour. Combined methodologies can reinforce the advantages while

mitigating the limitations of each of the monitoring techniques. Surveys, ethnographies, and direct observation give context to sensor readings that remain objective despite interim analyses. Using more than one monitoring method also increases the level of detail available. Sensors or spot-checks may give a picture of household characteristics, but surveys and structured observations can be used to inform individual behaviour. In turn, these combined sources provide further refinement for sensor algorithms or survey indices for streamlined analysis during subsequent monitoring periods. Comparisons and correlations across different monitoring methodologies can also be used to support the internal validity of results.

An important function of sensors is their ability to validate the reliability of another method while also suggesting improvements to standards of practice. For example, higher correlation between sensors and self-reports is seen when the questionnaire focuses on recent behaviour, particularly in the previous 48 hours, suggesting how surveys should be administered in future studies (Sinha et al., 2016). The appropriateness of structured observation as the gold standard is also questionable given reactivity that has been observed with passive latrine use monitors (Clasen et al., 2012). Similar findings were made with SmartSoap for handwashing behaviour (Ram et al., 2010). Integrated methods over appropriately sampled sub-populations could greatly reduce measurement and inferential uncertainty. For example, DelAgua Health Ltd's monitoring and evaluation programme linked household surveys and indicators, health outcomes, and sensors to determine the impact of a cook stove and water filter intervention in Rwanda (Thomas and Mattson, 2013; Rosa et al., 2014; Barstow et al., 2016; Nagel et al., 2016; Snoad et al., 2017). Although stronger monitoring methodologies and technologies can be cost-prohibitive at scale, continued cost reductions, widespread network coverage for mobile devices, and increased access to and familiarity with the Internet of Things and mobile devices could facilitate increased use of information and communication technologies for monitoring. When the budget, time, and training are available, combined methodologies can provide a more comprehensive and instructive depiction of WASH usage.

Conclusion

Although the SDG target for sanitation and hygiene represents a marked improvement over the MDG target, there is still a substantial disconnection between the desired improvements in service levels and the proposed indicators. The inclusion of safety, adequacy, and equity acknowledge that progress cannot be measured by simply counting the number of latrines or soap bars. Instead, health benefits from improved sanitation and hygiene facilities depend on the facilities' accessibility, their use, and their safety (their ability to effectively separate excreta from human contact both within and beyond the household). However, the proposed service ladders still rely heavily on direct observation of sanitation and handwashing facilities to infer usage, the safety of the facility, and the management of excreta. Similarly, household surveys depend on self-reported estimates of use, acceptability, and equity that can differ significantly from actual use due to courtesy

or recall bias. This combination of measurement error based on self-reporting and inferential error based on facility type can compound the uncertainty associated with estimates of real health impact due to improved sanitation and hygiene facilities. Technologies or methodologies that accurately measure sanitation outcomes – i.e. use, household safety, community safety, etc. – can reduce uncertainty by reducing measurement error and limiting the chain of causal inference between the indicator and the desired effect.

References

Arnold, B.F., Khush, R.S., Ramaswamy, P., Rajkumar, P., Durairaj, N., Ramaprabha, P., Balakrishnan K. and Colford, J.M. (2015) 'Short Report: Reactivity in rapidly collected hygiene and toilet spot check measurements: A cautionary note for longitudinal studies', *American Journal of Tropical Medicine and Hygiene* 92: 159–62.

Barreto, M.L., Genser, B., Strina, A., Teixeira, M.G., Assis, A.M., Rego, R.F., Teles, C.A., Prado, M.S., Matos, S.M., Santos, D.N., dos Santos, L.A. and Cairncross, S. (2007) 'Effect of city-wide sanitation programme on reduction in rate of childhood diarrhoea in northeast Brazil: assessment by two cohort studies', *Lancet* 370(9599): 1622–8.

Bangdiwala, S.I., Ramiro, L., Sadowski, L.S., Bordin, I.A., Hunter, W. and Shankar, V. (2004) 'Intimate partner violence and the role of socioeconomic indicators in WorldSAFE communities in Chile, Egypt, India and the Philippines', *Injury Control and Safety Promotion* 11(2): 101–9. doi:10.1080/15660970412331292324 [last accessed 15 August 2018].

Barstow, C.K., Nagel, C.L., Clasen, T.F. and Thomas, E.A. (2016) 'Process evaluation and assessment of use of a large scale water filter and cookstove program in Rwanda', *BMC Public Health* 16: 584 <https://doi.org/10.1186/s12889-016-3237-0> [last accessed 15th August 2018].

Biran, A., Rabie, T., Schmidt, W., Juvekar, S., Hirve, S. and Curtis V. (2008) 'Comparing the performance of indicators of hand-washing practices in rural Indian households', *Tropical Medicine and International Health* 13: 278–85.

Cairncross, S., Shordt, K., Zacharia, S. and Govindan, B.K. (2005) 'What causes sustainable changes in hygiene behaviour? A cross-sectional study from Kerala, India', *Social Science and Medicine* 61: 2212–20.

Cartwright, N. and Hardie, J. (2012) *Evidenced-Based Policy: A Practical Guide to Doing It Better,* New York Oxford University Press.

Clasen, T., Boisson, S., Routray, P., Torondel, B., Bell, M., Cumming, O., Ensink, J., Freeman, M., Jenkins, M., Odagiri, M., Ray, S., Sinha, A., Suar, M. and Schmidt, W.P. (2014) 'Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial', *The Lancet Global Health* 2: e645–53.

Clasen, T., Fabini, D., Boisson, S., Taneja, J., Song, J., Aichinger, E., Bui, A., Dadashi, S., Schmidt, W.P., Burt, Z. and Nelson, K.L. (2012) 'Making sanitation count: Developing and testing a device for assessing latrine use in low-income settings', *Environmental Science and Technology* 46: 3295–303.

Clasen, T. (2018) 'Improving safe water and sanitation monitoring for health gains', in *Innovations in WASH Impact Measures: Water and Sanitation Measurement Technologies and Practices to Inform the Sustainable Development Goals,* E. Thomas, L.E. Andrés, C. Borja-Vega, and G. Sturzenegger (eds), Directions in Development. Washington, DC: World Bank. <https://doi.org/10.1596/978-1-4648-1197-5> [last accessed 15th August 2018].

Coffey, D., Gupta, A., Spears, D., Khurana, N., Srivastav, N., Hathi, P. and Vyas, S. (2014) 'Revealed preference for open defecation', *Economic and Political Weekly* 49: 43–55.

Curtis, V.A., Danquah, L.O. and Aunger, R.V. (2009) 'Planned, motivated and habitual hygiene behaviour: An eleven country review', *Health Education Research* 24: 655–73.

Delea, M.G., Nagel C.L., Thomas E.A., Halder A.K., Amin N., Shoab, A.K., Freeman. M.C., Unicomb, L. and Clasen, T.F. (2017) 'Comparison of respondent-reported and sensor-recorded latrine utilization measures in rural Bangladesh: a cross-sectional study', *Transactions of the Royal Society of Tropical Medicine and Hygiene* 111(7): 308–315.

Escamilla, V., Knappett, P.S.K., Yunus, M., Streatfield, P.K. and Emch, M. (2013) 'Influence of latrine proximity and type on Tubewell water quality and diarrheal disease in Bangladesh', *Annals of the Association of American Geographers* 103(2): 299–308.

Feachem, R.G., D.J. Bradley, H. Garelick, and D.D. Mara. 1983. Sanitation and Disease: Health Aspects of Wastewater and Excreta Management. World Bank Studies in Water Supply and Sanitation. Washington, DC: World Bank.

Gadgil, M.A., Unicomb, L., Luby, S.P. and Ram, P.K. (2011) 'Consistent soap availability correlates with use of inexpensive soap products and improved handwashing in low-income households in Dhaka, Bangladesh', American Society of Tropical Medicine and Hygiene Annual Meeting, Philadelphia, PA, USA.

Geruso, M. and Spears, D. (2015) 'Neighborhood sanitation and infant mortality' (No. w21184), Washington, DC: National Bureau of Economic Research.

icddr,b (2008) 'Handwashing behavior in rural Bangladesh', *Health and Science Bulletin* 6(3): 2–7.

Jenkins, M.W. and Curtis, V. (2005) 'Achieving the "good life": Why some people want latrines in rural Benin', *Social Science and Medicine* 61: 2446–59.

Luby, S.P., Halder, A.K., Tronchet, C., Akhter, S., Bhuiya, A. and Johnston, R.B. (2009) 'Household characteristics associated with handwashing with soap in rural Bangladesh', *American Journal of Tropical Medicine and Hygiene* 81(5): 882–7.

Moraes, L.R., Cancio, J.A. and Cairncross, S. (2004) 'Impact of drainage and sewerage on intestinal nematode infections in poor urban areas in Salvador, Brazil', *Transactions of the Royal Society of Tropical Medicine and Hygiene* 98(4): 197–204.

Nagel, C.L., Kirby, M.A., Zambrano, L.D., Rosa, G., Barstow, C.K., Thomas, E.A. and Clasen, T.F. (2016) 'Study design of a cluster-randomized controlled trial to evaluate a large-scale distribution of cook stoves and water filters in Western Province, Rwanda', *Contemporary Clinical Trials Communications* 4: 124–35. <http://doi.org/10.1016/j.conctc.2016.07.003> [last accessed 15th August 2018].

Nique, M. and Smertnik, H. (2015) *Mobile for Development Utilities Programme,* GSMA: London. O'Reilly, K., Louis, E., Thomas, E. and Sinha, A. (2015) 'Combining sensor monitoring and ethnography to evaluate household latrine usage in rural India', *Journal of Water, Sanitation and Hygiene for Development* 5: 426–38.

Pickering, A.J., Blum, A.G., Breiman, R.F., Ram, P.K. and Davis, J. (2014) 'Video surveillance captures student hand hygiene behavior, reactivity to observation, and peer influence in Kenyan primary schools', *PLoS ONE* 9: 1–7.

Pickering, A.J., Boehm, A.B., Mwanjali, M. and Davis, J. (2010) 'Efficacy of waterless hand hygiene compared with handwashing with soap: a field study in Dar es Salaam, Tanzania', *American Journal of Tropical Medicine and Hygiene* 82(2): 270–8.

Ram, P. (2013) 'Practical guidance for measuring handwashing behavior: 2013 Update', Global Scaling Up Hand-washing Project, Water and Sanitation Project, The World Bank <https://openknowledge.worldbank.org/handle/10986/19005> [last accessed 15th August 2018].

Ram, P.K., Halder, A.K., Granger, S.P., Jones, T., Hall, P., Hitchcock, D., Wright, R., Nygren, B., Islam, M.S., Molyneaux, J.W. and Luby, S.P. (2010) 'Is structured observation a valid technique to measure handwashing behavior? Use of acceleration sensors embedded in soap to assess reactivity to structured observation', *American Journal of Tropical Medicine and Hygiene* 83: 1070–6.

Ram, P.K., Jahid, I., Halder, A.K., Nygren, B., Islam, M.S., Granger, S.P., Molyneaux, J.W. and Luby, S.P. (2011) 'Variability in hand contamination based on serial measurements: implications for assessment of hand-cleansing behavior and disease risk', *American Journal of Tropical Medicine and Hygiene* 84(4): 510–6.

Rand, E.C., Loughnan, L.C., Maule, L. and Reese, H.E. (2015) 'Management of child feces: current disposal practices', Water and Sanitation Program: Research Brief (June), p. 8, available from: <https://www.wsp.org/sites/wsp.org/files/publications/WSP-CFD-Summary-Brief.pdf> [last accessed 15th August 2018].

Ribot, J.C. and Peluso, N.L. (2003) 'A Theory of Access', *Rural Sociology* 68: 153–81.

Robiarto, A., Sofyan, E., Setiawan, D., Malina, A. and Rand, E. (2014) 'Scaling up Indonesia's rural sanitation mobile monitoring system nationally', Water and Sanitation Program learning note, International Bank for Reconstruction and Development/The World Bank. <https://www.wsp.org/sites/wsp.org/files/publications/WSP-Indonesia-Mobile-Monitoring.pdf> [last accessed Accessed 15 August 2018].

Rosa, G., Majorin, F., Boisson, S., Barstow, C., Johnson, M., Kirby, M., Ngabo, F., Thomas, E. and Clasen, T. (2014) 'Assessing the impact of water filters and improved cook stoves on drinking water quality and household air pollution: a randomised controlled trial in Rwanda', *PLoS ONE* 9(3): 1–9.

Sagerman, D., Nizame, F., Yu, J., Luby, S.P. and Ram, P. (2011) 'Impact of complexity of instructions on handwashing adherence in a low income setting', the American Society of Tropical Medical and Hygiene conference Dhaka, Bangladesh.

Sanipath (2014) Rapid Assessment Tool manual draft, 7 October sanipath.org [last accessed 15th August 2018].

Schwemlein, S., Cronk, R. and Bartram, J. (2016) 'Indicators for monitoring water, sanitation, and hygiene: A systematic review of indicator selection methods', *International Journal of Environmental Research and Public Health* 13(3).

Sclar G.D., Penakalapati, G., Caruso, B., Rehfuess, E., Garn, J., Alexander, K., Freeman, M., Boisson, S., Meddlicot, K. and Clasen, T. (2018) 'Exploring the relationship beween sanitation and well-being: a systematic review and best-fit framework synthesis', article forthcoming.

SHARE (2014) *Sanitation Vulnerability: Women's Stress and Struggles for Violence-free Sanitation,* London, UK: London School of Hygiene and Tropical Medicine.

Shuval, H. (2003) 'Estimating the global burden of thalassogenic diseases: human infectious diseases caused by wastewater pollution of the marine environment', *Journal of Water Health* 1(2): 53–64.

Sinha, A., Nagel, C.L., Thomas, E., Schmidt, W.P., Torondel, B., Boisson, S. and Clasen, T.F. (2016) 'Assessing latrine use in rural India: a cross-sectional study comparing reported use and passive latrine use monitors', *American Journal of Tropical Medicine and Hygiene*.

Snoad, C., Nagel, C., Bhatacharya, A. and Thomas, E. (2017) 'The effectiveness of sanitary inspections as a risk assessment tool for theromotolerant coliform bacteria contamination of rural drinking water: A review of data from West Bengal India,' *The American Journal of Tropical Medicine and Hygiene* 96(4): 976–83.

Thomas, E.A. and Mattson, K. (2013) 'Instrumented monitoring with traditional public health evaluation methods: An application to a water, sanitation and hygiene program in Jakarta', Mercy Corps technical report <https://www.mercycorps.org/sites/default/files/Instrumented%20Monitoring%20Indonesia.pdf> [last accessed Accessed 15 August 2018].

Webb, A.L., Stein, A.D., Ramakrishnan, U., Hertzberg, V.S., Urizar, M. and Martorell, R. (2006) 'A simple index to measure hygiene behaviours', *International Journal of Epidemiology* 35: 1469–77.

WHO/UNICEF (2015) JMP Green Paper, 'Global monitoring of water, sanitation and hygiene post-2015', World Health Organization: Geneva.

WSP/UNICEF (2015) 'Child Feces Disposal in India', available from: <http://www.wsp.org/sites/wsp.org/files/publications/WSP-India-CFD-Profile.pdf> [last accessed 15th August 2018].