Received: 10 April 2021 | Revised: 16 September 2021 | Accepted: 18 November 2021

DOI: 10.1002/hec.4462

Health Economics WILEY

RESEARCH ARTICLE

Measuring and valuing broader impacts in public health: Development of a sanitation-related quality of life instrument in Maputo, Mozambique

Ian Ross[1,2] | **Giulia Greco**[1,3] | **Charles Opondo**[4,5] | **Zaida Adriano**[6] | **Rassul Nala**[7] | **Joe Brown**[8] | **Robert Dreibelbis**[2] | **Oliver Cumming**[2]

[1]Department of Global Health and Development, London School of Hygiene and Tropical Medicine, London, UK

[2]Department of Disease Control, London School of Hygiene and Tropical Medicine, London, UK

[3]MRC/UVRI & LSHTM Uganda Research Unit, Kampala, Uganda

[4]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK

[5]Nuffield Department of Population Health, University of Oxford, Oxford, UK

[6]WE Consult, Maputo, Mozambique

[7]Instituto Nacional de Saúde, Maputo, Mozambique

[8]Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill, North Carolina, USA

Correspondence
Ian Ross, Department of Global Health and Development, London School of Hygiene and Tropical Medicine, LSHTM, 15-17 Tavistock Place, London, WC1H 9SH, UK.
Email: ian.ross@lshtm.ac.uk

Funding information
Economic and Social Research Council
Bill and Melinda Gates Foundation

Abstract

Two billion people globally lack access to a basic toilet. While improving sanitation reduces infectious disease, toilet users often identify privacy, safety and dignity as more important. However, these outcomes have not been incorporated in sanitation-related economic evaluations. This illustrates the general challenge of outcome measurement and valuation in the economic evaluation of public health interventions, and risks misallocating the US$ 20 billion invested in sanitation in low- and middle-income countries every year. In this study in urban Mozambique, we develop an instrument to measure sanitation-related quality of life (SanQoL). Applying methods from health economics and the capability approach, we develop a descriptive system to measure five attributes identified in prior qualitative research: disgust, health, shame, safety and privacy. Sampling individuals from the intervention and control groups of a sanitation intervention trial, we elicit attribute ranks to value a SanQoL index and assess its validity and reliability. In combination with a measure of time using a sanitation service, SanQoL can quantify incremental benefits in a sanitation-focused cost-effectiveness analysis. After monetary valuation based on willingness to pay, QoL benefits could be summed with health gains in cost-benefit analysis, the most common method in sanitation economic evaluations.

KEYWORDS
environment and development, environmental health, health behaviour, public health, quality of life, sanitation

1 | INTRODUCTION

Nearly two billion people globally lack access to basic sanitation (WHO & UNICEF, 2021). This failure to separate human excreta from human contact represents an enormous public health challenge. Every year 26 million disability-adjusted life years (DALYs), including 430,000 child deaths from diarrhoea, are attributable to inadequate sanitation (Prüss-Ustün et al., 2019). From a health system perspective, averting this disease burden would be highly valuable. Toilet users, however, often identify benefits of improved sanitation for privacy, safety, reduced smells, and dignity, as more important than reducing infectious disease (Elmendorf & Buckles, 1980; Jenkins & Curtis, 2005; Mukherjee, 2001). These are attributes of good quality of life (QoL), and contribute to health in its broadest sense including mental and social wellbeing (WHO, 1948). Therefore, the QoL benefits beyond infectious disease are likely to underpin household willingness to pay (WTP) for sanitation improvements, with consequences for their measurement and valuation in cost-benefit analysis (CBA).

Studies have identified associations between better sanitation and general measures of mental wellbeing (Caruso et al., 2018; Gruebner et al., 2012). However, the QoL outcomes such as privacy and safety are not captured in DALYs. Quality-adjusted life years (QALYs) are less-used in lower- and middle-income countries (LMICs), and have never been used to evaluate sanitation interventions. In any case, QALYs would also be unlikely to capture benefits of sanitation beyond morbidity and mortality from infectious disease. This is because changes in the QoL outcomes prioritised by toilet users would not be picked up by measures widely used to weight QALYs such as the EQ-5D, with the exception of consequences for anxiety/depression (Brazier et al., 2016; Euroqol Group, 2009).

Despite the importance of outcomes such as privacy and safety to users, economic evaluations of sanitation interventions have never managed to specifically include them, in the absence of means for their measurement and valuation (Hutton & Chase, 2016). The only way in which sanitation-related QoL outcomes have been valued economically is through WTP for toilets (Hutton & Chase, 2016). Willingness to pay for benefit valuation is often considered dissatisfactory in a health context due to the biases it introduces (Cookson, 2003), particularly from ability to pay (Coast, Smith, & Lorgelly, 2008). The bias is likely to be toward under-valuation if a respondent's WTP for a toilet is taken to comprise the total benefits of sanitation (e.g., morbidity, mortality, healthcare savings, time savings, privacy, safety, etc.). Accurate estimation is more likely if these benefits are estimated separately and aggregated, as is preferred in CBA (Boardman et al., 2018), and in practice all sanitation CBA studies do this (Hutton & Chase, 2016).

No quantitative measure capturing sanitation interventions' QoL outcomes with a user-derived valuation scheme exists (Sclar et al., 2018). The only standalone measure of such outcomes is the women's sanitation insecurity profile, which comprises 60 survey questions spread over seven factor scales measuring a mixture of practices, experiences and feelings (Caruso et al., 2017). While it is appropriate for non-economic purposes, the women's sanitation insecurity profile's application of equal weighting and lack of an overall score preclude its use in economic evaluation (Brazier et al., 2016). It was also developed with (and designed for) women only, while sanitation affects the general population. To inform the allocation of public funds, a respondent-weighted index with a single overall score designed for use in the general population is needed.

The lack of an outcome measure applicable for use in economic evaluation may be contributing to misallocation of the US$ 20 billion invested annually in sanitation in LMICs (WHO, 2017). Health economists are increasingly using capability-based outcome measures as a way to broaden the evaluative space beyond the value of health alone (Coast, Smith, & Lorgelly, 2008; Greco et al., 2016). The capability approach focuses on the value of what people are able to be and do (Sen, 1980, 1993) and is well-suited to a multi-dimensional conceptualisation of QoL (Coast et al., 2015). Many capability-based measures have been developed for different economic evaluation purposes both in high-income and low-income settings (Coast, 2019; Greco, 2016; Simon et al., 2013).

In Mozambique, where our study took place, 39% of the urban population (77% in rural) do not have basic sanitation, defined as “improved” facilities which are not shared with other households (WHO & UNICEF, 2021). A third of those use a shared improved facility. Urban sanitation services are a municipal responsibility and each city has policies and

strategies in place (CMM, 2017; MOPH, 2011), but access to basic sanitation areas is increasing at only 1.5 percentage points per year (WHO & UNICEF, 2021).

In this study, we aim to develop and value a measure of sanitation-related quality of life (SanQoL) in urban Mozambique, based on prior qualitative research in the setting (Ross et al., 2021). The underlying objective was to enable classification of "sanitation states" and their valuation, in support of economic evaluations of sanitation interventions. We include exploration of the measure's validity and reliability. The study was nested within the Maputo Sanitation (MapSan) trial (clinicaltrials.gov registration: NCT02362932), which evaluated the impact of a shared urban sanitation intervention on children's enteric infections (Knee et al., 2021).

2 | METHODS

2.1 | Study setting

We undertook this study in Maputo, Mozambique, in the low-income neighbourhoods of the Nhlamankulu district where multi-household compounds with shared sanitation facilities are common (Brown et al., 2015). Of Maputo City's 1.1 million population, 70% live in informal settlements (INE, 2019; UN-HABITAT, 2010). Approximately 89% use non-sewered sanitation facilities (Hawkins & Muximpua, 2015). In the control group of the MapSan trial, households used low-quality shared pit latrines, which are common in low-income areas. These latrines typically comprised unlined pits with squatting slabs made of wood or car tyres, and no water seal (u-bend) to provide a barrier against smells and flies. Few pit latrines have roofs, and the walls are often made with sections of scrap corrugated iron or plastic sheeting, with makeshift fabric doors. Together, these conditions lead to unpleasant, unsafe toilets providing little privacy.

In the intervention group, an international non-government organisation provided a subsidised pour-flush toilet with concrete superstructure, discharging to a septic tank. The intervention toilets had two design types, depending on user numbers. The first was a shared toilet with one stance (cubicle) to be used by around 15 people, at 85% subsidy. The second was a Community Sanitation Block with two stances, to be used by a minimum of 21 people, at 90% subsidy. Both STs and CSBs were built from concrete blocks, with metal doors lockable from the inside. More information on the setting, including photos of toilet types, is provided in Supplementary Material A. Our study received approval from the *Comité Nacional de Bioética para a Saúde* (IRB00002657) at the Ministry of Health in Mozambique and the ethics committee at the London School of Hygiene & Tropical Medicine (Ref: 14609). Informed, written consent was obtained from all participants. Deidentified individual participant data, codebook, and replication code are available online (Ross, 2021).

2.2 | Overall study design

We followed measure development methods common in health economics for the purpose of economic evaluation (Goodwin & Green, 2016), whereby the final measure comprises index values anchored at 0 and 1 (Drummond et al., 2015). The descriptive system comprises a set of items with categorical response scales, each representing one dimension of the construct being measured (Brazier et al., 2016). Items are selected primarily for content validity (Fayers & Machin, 2015), and an individual's combination of responses under the descriptive system comprises their "health state" (Brazier et al., 2016) or "capability state" (Coast, Flynn, et al., 2008). Methods such as factor analysis are not applied, and instead states are valued as an index using preference elicitation.

We adapted these methods to develop a measure of SanQoL. The capability approach to welfare economics (Sen, 1980, 1993) comprised our overarching theoretical approach to defining and measuring QoL, which informed both qualitative and quantitative methods. The target population for our measure is people living in urban settings where poor sanitation is common. The primary intended use is in economic evaluation of sanitation programmes. This required the measure to have a small number of items, since trading-off more than seven attribute levels simultaneously in preference elicitation tasks is generally considered too cognitively demanding (Hensher et al., 2015; Mangham et al., 2009).

Initial qualitative research in the same Maputo setting defined the construct to be measured as "the subset of overall QoL which is directly affected by sanitation practices or services" (Ross et al., 2021). This draws on common definitions of health-related QoL (HRQoL), and is aligned with the third of the four types identified by Karimi and Brazier (2016) in their review of definitions of health and HRQoL. Specifically we build on how Peasgood et al. (2014, p. 17) describe narrower definitions of HRQoL: "the sub-set of the important or most common ways in which health or health care impact

upon well-being." The scope of sanitation practices is as perceived by users but is assumed to extend beyond defecation and urination to include for example menstrual hygiene, as well as any related practices users consider important, such as bathing. The capability approach was reflected in the design of the qualitative study (Ross et al., 2021). For example, the topic guides focused on a broad evaluative space ("a good life"), we used focus groups to engender the deliberation encouraged by the capability approach, and we triangulated findings on attributes' relative importance using methods from cognitive anthropology (Weller & Romney, 1988).

The qualitative research resulted in a conceptual model (Figure 1) centred on five attributes: disgust, health, privacy, safety and shame. These were the attributes toilet users in this setting considered most important, and are largely consistent with the broader qualitative literature in other countries and settings (Novotný et al., 2018). Some studies have identified convenience or water supply, for example, as important (Jenkins & Curtis, 2005; Sahoo et al., 2015). However, these were not raised by our participants, reflecting that the social and environmental context in which a toilet "commodity" (Sen, 1985) is used affects the capabilities an individual can derive from it (Ross et al., 2021).

In this paper, we present the subsequent stages of measure development (Table 1), adhering to the standards of the International Society for QoL Research for patient-reported outcome measures (Supplementary Material B) (Reeve et al., 2013). We conducted all analyses in Stata 16 (StataCorp, 2019).

2.3 | Item development and piloting

We developed several possible items for each of the five attributes identified by the qualitative research (Streiner et al., 2015), aiming to focus attention on capabilities by framing questions with "can you...?" or "are you able to...?" A long-list of 36 items was reviewed internally, then by 14 external experts. We undertook piloting and cognitive interviewing during April 2019, recruiting a team of four enumerators fluent in Portuguese and Changana, the first and second most commonly spoken languages in Maputo. One author (Z. A.), a Mozambican fluent in Portuguese and English, translated items into Portuguese. (I. R.), an intermediate Portuguese speaker, extensively discussed meanings with (Z. A.) and the enumerator team. We piloted items with 64 individuals from the target population, and undertook further cognitive interviewing with 28 of the pilot respondents (Bowden et al., 2002). In cognitive interviews, after each SanQoL item, the respondent was asked to explain back the question in their own words and discuss the ease of understanding the question. Enumerators rated their explanation on a scale of full, partial or no comprehension. For each of the five attributes, we identified one item which best achieved face and content validity, considering the qualitative findings and cognitive interviews. These five items comprise the descriptive system, with responses on a four-level ordinal frequency scale: always, sometimes, rarely, never (Table 2, and in Portuguese in Supplementary Material C). The descriptive system

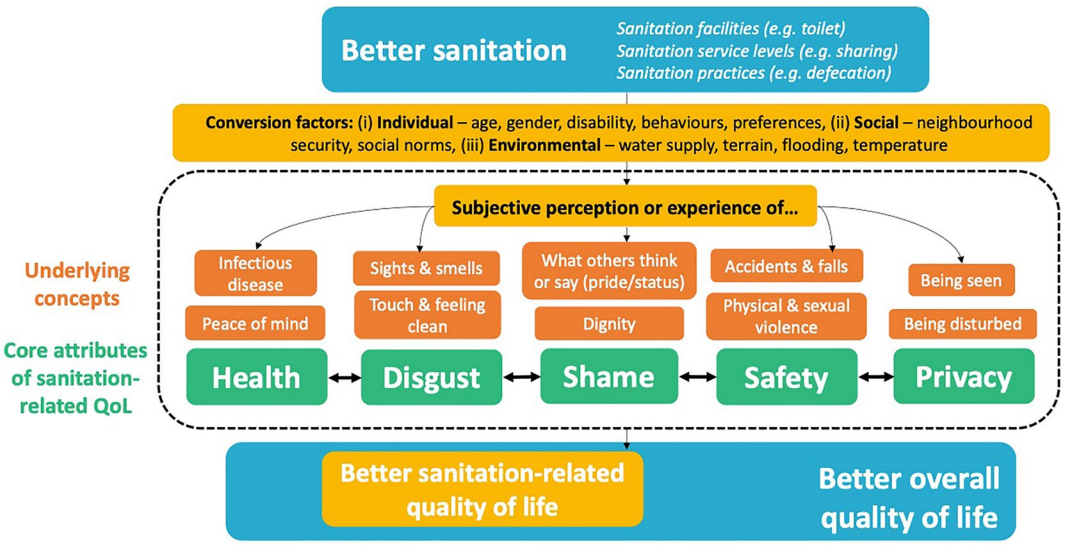

FIGURE 1 Conceptual model for sanitation-related quality of life (Ross et al., 2021)

TABLE 1 Stages followed in measure development

Stage	Objective	Method	Main output
1. Conceptual model	To identify what people most value about sanitation, using qualitative research	In-depth interviews ($n = 19$) and focus groups ($n = 8$)	Conceptual model
2. Piloting	To ensure items are understood, and identify floor/ceiling effects	Pilot interviews ($n = 64$) and cognitive interviews ($n = 28$)	Five-item descriptive system
3. Valuation	To estimate the relative value of attributes	Rank sum method ($n = 424$) with a view to subsequent discrete choice experiment	SanQoL index values anchored at 0 and 1
4. Validity & reliability	To assess the extent to which the instrument measures what it intends to measure (validity) and does so consistently (reliability)	Quantitative survey amongst intervention and control groups of a trial ($n = 424$)	Evidence of construct validity, convergent validity, and test-retest reliability.

Abbreviation: SanQOL, sanitation-related quality of life.

TABLE 2 Descriptive system for the SanQoL measure

Attribute	Questionnaire item	Responses[a]
Disgust	Can you use the toilet without feeling disgusted?	• Always • Sometimes • Rarely • Never
Health	Can you use the toilet without worrying that it spreads diseases?	
Privacy	Can you use the toilet in private, without being seen?	
Shame	Can you use the toilet without feeling ashamed for any reason?	
Safety	Are you able to feel safe while using the toilet?	

Abbreviation: SanQoL, sanitation-related quality of life.

[a]respondents can choose "prefer not to answer" for any item. The recall period was 4 weeks (Supplementary Material C).

therefore contains 1024 ($=4^5$) combinations of attribute levels, or "sanitation capability states." Items were framed as direct questions such that "always" was consistently the best outcome.

2.4 | Valuation

Valuation aims to aggregate item responses into a single score, weighted by the relative value of attributes elicited from the target population. We undertook a quantitative survey in May 2019 with the same enumerator team, using the mWater (2019) application on smartphones. We aimed to recruit at least 400 people aged over 18 living on trial-enrolled compounds, stratified by intervention status and gender. This meant a range of outcomes was likely, due to the diversity of low-quality toilet types in control compounds and the high-quality toilets in use in intervention compounds. We recruited a maximum of two people (one man, one woman) from each compound, on condition that they were not from the same household. In addition to SanQoL items, we collected data on water supply and sanitation services, demographic characteristics, and calculated an asset-based wealth index following standard methods (Vyas & Kumaranayake, 2006). Finally, we asked respondents to rank the five SanQoL attributes on a visual analogue scale, according to their relative importance (Supplementary Material D).

With a view to subsequent valuation by discrete choice experiment, our interim valuation strategy was based on the rank sum method (Stillwell et al., 1981), following previous measures (de Kruijk & Rutten, 2007; Greco, 2016). This required scoring attributes on a 0-3 scale based on item responses (0 = never, 1 = rarely, 2 = sometimes, 3 = always), and estimating attribute weights from the rank data (Equation 1). We calculated SanQoL index values by combining attribute scores and weights, then rescaling to anchor at 0 and 1 (Equation 2). Zero represents "no sanitation capability" and one represents "full sanitation capability," building on the Investigating Choice Experiments and Capabilities measures' anchor points (Coast, Flynn, et al., 2008).

Equation 1 – attribute weights for a population:

$$w_i = \frac{N - R_i + 1}{\sum_{i=1}^{N}\left(N - R_i + 1\right)} \tag{1}$$

Equation 2 – SanQoL index value for an individual:

$$S_j = \frac{\sum_{i=1}^{N} \left(x_{ij} * \mathrm{w}_i \right)}{3} \tag{2}$$

where: w_i is the weight of the ith attribute; N is the number of attributes; R_i is the mean rank of the ith attribute in the population; x_{ij} are attribute scores ranging from 0 to 3 for the jth individual, where "always" = 3 and "never" = 0; S_j is the SanQoL index value for the jth individual.

We explored differences in attribute ranks by gender, whether the respondent was elderly (aged 60 or over), and treatment group. The rationale for including gender was that women experience sanitation differently to men in that they squat for urination, manage menstrual hygiene, and are more likely to fear and experience "peeping" or assault (Tilley et al., 2013). The rationale for including elderliness was that older people are particularly impacted by poor sanitation, predominantly as a result of disabilities that occur with ageing (Groce et al., 2011). In Mozambique, an elderly person is defined in law as anyone aged 60 or older (Castel-Branco & Andrés, 2019). The rationale for including treatment group was to assess whether people who had been using higher-quality toilets for four years (intervention group) had different perceptions of the relative value of the attributes. Rank is an ordered categorical variable, so we used mixed-effects ordered logit models, clustering standard errors at the compound level. We regressed on the rank for each attribute (which ranged from 1 to 5), including as covariates gender, aged 60 or over, and treatment. We also included as covariates any variables which both (i) differed at the 5% level along any of those three lines; and (ii) plausibly influenced ranking. The two variables meeting those criteria were the wealth index (differed by treatment status) and primary education completion (differed by gender and by aged 60 or over). *P*-values less than 0.05 were considered statistically significant evidence of association. As a robustness check, we also used generalised linear mixed models (GLMM) to analyse rank as a continuous variable, with the same set of covariates.

2.5 | Validity and reliability assessment

We assessed internal reliability using item-total correlation and Cronbach's alpha, for which common acceptability thresholds are 0.4 (Ware et al., 1980) and 0.7 (Nunnally, 1978) respectively. There is debate about whether these metrics are appropriate for indices, since each item measures a different dimension (Konerding, 2013), but we include them because our measure comprises a single overall score to represent a single construct. We examined distributions of frequency endorsements in aggregate and by gender (Terwee et al., 2007). We assessed test-retest reliability by re-interviewing 69 respondents (16% of sample) two weeks after the original interview (Streiner et al., 2015). We used a two-way mixed effects model (Koo & Li, 2016) to evaluate the intraclass correlation coefficient (ICC) for the SanQoL index value, against an acceptability threshold of 0.7 (Terwee et al., 2007).

To assess construct validity, we pre-specified hypotheses (Table 3) about the presence or absence and direction of associations between each attribute score and a set of user and toilet characteristics. We tested these using GLMM, because attribute scores represent points on an underlying continuous scale, with standard errors clustered at the compound level. We regressed on each attribute score, including as covariates: gender; aged 60 or over; wealth index; primary education completion; whether the toilet had a concrete or tile floor; whether the toilet had masonry or zinc sheet walls; whether the toilet had an inside lock; and, whether the enumerator smelt faeces on entering toilet. The hypotheses drew on the previous qualitative work (Ross et al., 2021) and the broader literature on motives for sanitation behaviours and mental wellbeing (Novotný et al., 2018; Sclar et al., 2018).

We assessed convergent validity by correlation (Pearson's r) between SanQoL index values and the WHO-5 mental wellbeing index (Topp et al., 2015). We expected correlation to be positive but less than 0.5, since sanitation is unlikely to be a primary driver of mental wellbeing (Sclar et al., 2018). Finally, we investigated the convergence of SanQoL index values between respondents using the same toilet. We used inter-rater methods to test the hypothesis that responses would be positively correlated but not equal, because any two people may experience the same toilet differently. We calculated the ICC using a one-way random effects model (Koo & Li, 2016). Interpretation of this ICC is "fair" (0.40–0.59), "good" (0.60–0.74), or "excellent" (>0.75) (Cicchetti, 1994).

TABLE 3 Hypothesised presence or absence of associations

	Association hypothesised					
Variable	**Disgust**	**Health**	**Shame**	**Safety**	**Privacy**	**Rationale for hypothesis**
User characteristics						
Female	No	No	No	Yes (−ve)	Yes (−ve)	Women might have higher acceptability thresholds for safety and privacy, since they are at higher risk of peeping, sexual harassment, and assault.
Aged 60+	No	No	No	No	No	No reason to expect any individual item to systematically covary with being elderly.
Wealth index (continuous)	No	No	No	No	No	No reason to expect any individual item to systematically covary with wealth.
Toilet characteristics						
Toilet floor material (high/low quality)	Yes (+ve)	Yes (+ve)	Yes (+ve)	Yes (+ve)	No	The quality of the toilet floor might affect all attributes, but there is no obvious rationale for privacy.
Toilet wall material (high/low quality)	No	No	Yes (+ve)	Yes (+ve)	Yes (+ve)	The quality of the toilet wall might affect the extent to which it prevents others from seeing toilet users, thereby affecting privacy, safety and shame.
Toilet locks from the inside	No	No	No	Yes (+ve)	Yes (+ve)	An inside lock might directly improve privacy and safety, but there is no obvious rationale for disgust, health and shame.
Enumerator smells faeces on entry	Yes (−ve)	Yes (−ve)	Yes (−ve)	No	No	A smelly toilet is likely to affect disgust, shame and perception of health risk. There is no obvious rationale for safety and privacy.

Note: "Yes (−ve)" means that having the specified characteristic is hypothesised to be associated with lower (worse) scores for the attribute, for example, women are hypothesised to have lower privacy scores than men, all else being equal. "Yes (+ve)" indicates a hypothesis of higher scores and "no" of no significant association.

3 | RESULTS

3.1 | Piloting and cognitive interviews

Of the 64 piloting respondents, 48% were male and 52% female. No single category per item received greater than 50% of endorsements in piloting, well-below an 80% benchmark (Streiner et al., 2015). As is desirable, there were no floor or ceiling effects, that is, absence of high proportion of respondents with maximum/minimum scores (Streiner et al., 2015). Cognitive interviewing showed that items could be understood, and no changes to the English descriptive system were required. However, there was one change to the Portuguese. The disgust item (Table 2), originally framed as an adjective (*enojado* = disgusted) was reframed as a noun (*nojo* = disgust), considered more natural in the Portuguese spoken in the setting.

3.2 | Valuation

We interviewed 424 individuals (220 female, 204 male) each from a different household, with a response rate of 99%. Respondent characteristics are tabulated by gender in Table 4 (and by trial arm in Supplementary Material E). Respondents lived on 275 MapSan trial compounds (131 control, 144 intervention). About two thirds had completed primary education, with slightly more men (70%) than women (57%) having done so. There was near-universal access to piped water connections (98%). The vast majority (89%) of respondents shared their toilet with other households, with a mean of 12.2 people sharing each toilet stance (cubicle).

Health had the highest mean rank and shame the lowest, meaning they also had the highest and lowest attribute weights (Table 5). Using Equation 2, we derived a value set (Supplementary Material E). Across the five attributes, there was insufficient evidence for differences in ranks by gender, being elderly, or intervention status (Table 6, full output in

TABLE 4 Characteristics of quantitative sample, overall and by gender

	Overall ($n = 424$)	Male ($n = 204$)	Female ($n = 220$)
Respondent demographic characteristics			
Respondent age	39.9 (15.3)	39.3 (15.1)	40.5 (15.5)
Respondent has a partner	214 (50%)	116 (57%)	98 (45%)
Household size	5.1 (3.0)	4.7 (3.0)	5.5 (2.9)
Number of children under-14	1.3 (1.6)	1.1 (1.3)	1.4 (1.7)
Other respondent characteristics			
Completed primary school or above	268 (63%)	143 (70%)	125 (57%)
Completed secondary school or above	51 (12%)	30 (15%)	21 (10%)
Moderate problems walking about, or worse	25 (6%)	9 (4%)	16 (7%)
Moderate pain or discomfort, or worse	38 (9%)	12 (6%)	26 (12%)
Respondent housing			
Dwelling has cement or tiled floor	394 (93%)	191 (94%)	203 (92%)
Dwelling has concrete exterior walls	283 (67%)	131 (64%)	152 (69%)
Dwelling has zinc or concrete roof	424 (100%)	204 (100%)	220 (100%)
Household has access to electricity connection	359 (85%)	175 (86%)	184 (84%)
Rents dwelling	114 (27%)	56 (27%)	58 (26%)
Compound-level water and sanitation characteristics			
Piped water connection	416 (98%)	199 (98%)	217 (99%)
Water available at least 8 h/day	209 (49%)	101 (50%)	108 (49%)
Uses on-plot toilet	416 (98%)	201 (99%)	215 (98%)
Pour-flush to septic tank (intervention)	222 (52%)	103 (51%)	119 (54%)
Pit latrine (control)	202 (48%)	101 (49%)	101 (46%)
Shares toilet with other household(s)	377 (89%)	186 (91%)	191 (87%)
Number of households sharing stance	3.2 (1.6)	3.3 (1.6)	3.2 (1.7)
Number of people sharing stance	12.2 (6.0)	12.2 (6.2)	12.2 (5.8)

Note: Data are *n* (%) for categorical variables and mean (SD) for numerical variables. For categorical variables (e.g., respondent has a partner) the number of people which have that characteristic is presented. In the publicly available dataset we categorised age, household size and the number of children under 14 to maintain full anonymity, since several values were shared by 5 people or fewer. This table reports the mean of the continuous versions of those variables.

TABLE 5 Mean ranks of SanQoL attributes from ranking exercise, with calculated weights

	Mean rank						
	Overall		Female		Male		Weight for index valuation (from "overall")
	Mean	SD	Mean	SD	Mean	SD	
Disgust	2.74	1.29	2.72	1.31	2.77	1.26	0.22
Health	1.72	0.96	1.72	0.94	1.72	0.98	0.29
Shame	4.06	1.15	4.13	1.03	4.00	1.27	0.13
Safety	3.50	1.30	3.53	1.34	3.46	1.25	0.17
Privacy	2.96	1.18	2.89	1.16	3.04	1.21	0.20

Abbreviation: SanQoL, sanitation-related quality of life.

Supplementary Material E), which is broadly desirable as heterogeneity in public values can make decision-making more contested (Baker et al., 2021). The GLMM robustness checks support this result (Supplementary Material E).

Control group respondents used pit latrines of various levels of quality, and their SanQoL index values approximate a normal distribution (Figure 2). The distribution for intervention group respondents is unimodal but left-skewed because toilets were high-quality. The mean SanQoL index value was 0.49 in the control group sample (95% CI: 0.46–0.53) and 0.83 in the intervention group (95% CI: 0.81–0.86).

TABLE 6 Evidence for difference in attribute ranks by user characteristics (mixed effects ordered logit)

		Disgust	Health	Shame	Safety	Privacy
Female	Odds ratio (95% CI)	0.16 (0.64–1.30)	0.22 (0.72–1.60)	0.28 (0.80–1.93)	0.20 (0.74–1.56)	0.15 (0.55–1.15)
	p-value	0.62	0.72	0.32	0.71	0.22
Aged 60+	Odds ratio (95% CI)	1.21 (0.69–2.10)	0.51* (0.24–1.09)	1.03 (0.56–1.90)	1.15 (0.64–2.06)	0.79 (0.42–1.50)
	p-value	0.51	0.08	0.91	0.65	0.47
Treatment	Odds ratio (95% CI)	0.70* (0.47–1.04)	1.08 (0.68–1.73)	1.08 (0.66–1.76)	1.36 (0.90–2.05)	0.86 (0.56–1.32)
	p-value	0.08	0.74	0.76	0.15	0.49

Note: Standard errors are clustered at the compound level. *, **, *** indicate significance at the 10, 5 and 1 percent level.

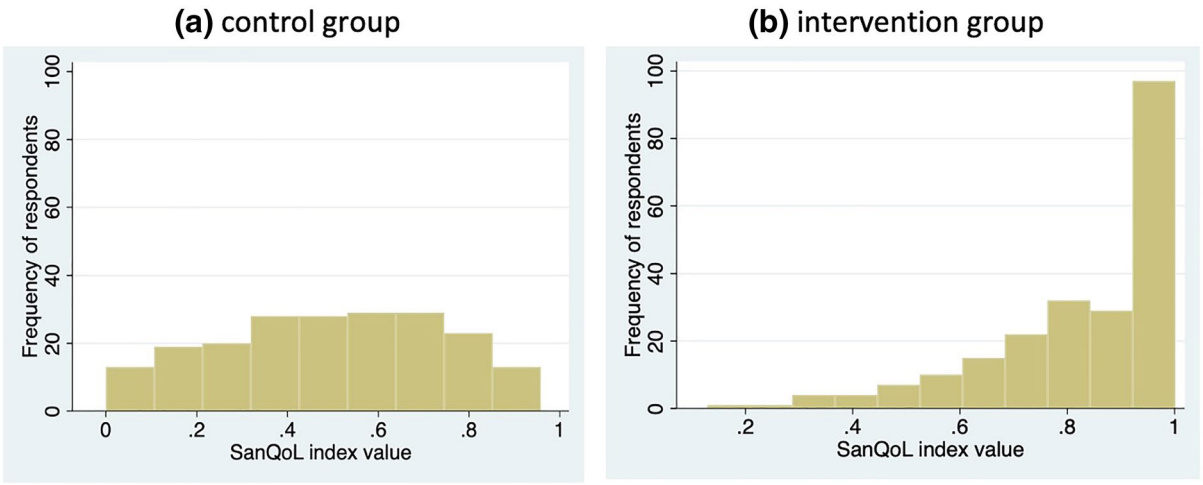

FIGURE 2 Histograms of sanitation-related quality of life (SanQoL) index values by treatment group

3.3 | Validity and reliability assessment

There were no missing values for SanQoL attributes. For the privacy item, 61% of women and 57% of men endorsed "always," which is slightly elevated but not enough to cause concern in this sample (Table 7). There were only eight participants who responded "never" to all five questions, and a further five who responded "never" for all but one of the five questions. There were 75 participants who responded "always" to all five questions.

Item-total correlations ranged between 0.69 and 0.73 (Supplementary Material F). Cronbach's alpha was 0.77 and the test-retest ICC for the index was 0.87, both above specified thresholds. Inter-item polychoric correlation coefficients averaged 0.51, with range 0.40–0.70, supporting the idea that attributes are sufficiently distinct but nonetheless measure the same construct (Supplementary Material F).

Considering construct validity, there was evidence of association for 86% (30/35) of hypothesised presence or absence associations, or 79% (11/14) considering presence only (Table 8 and Supplementary Material G). All presence associations were in the hypothesised direction. Correlation between SanQoL index values and the WHO-5 mental wellbeing index was 0.24, indicating slight convergence as hypothesised. The ICC was 0.72 for convergence of SanQoL index values between two respondents using the same toilet, indicating substantial but not complete correlation, as hypothesised.

4 | DISCUSSION

In this study, we developed a SanQoL instrument in urban Mozambique, with a five-item descriptive system and accompanying valuation. By sampling individuals from the intervention and control groups of a sanitation intervention trial, we undertook an interim valuation of the SanQoL index, and assessed its validity and reliability. Considering construct validity, there was evidence supporting 79% of hypothesised presence associations, all in the hypothesised direction.

TABLE 7 Endorsement frequencies for SanQoL items, by gender

Response	Overall ($n = 424$) *N* (%)	Female ($n = 220$) *N* (%)	Male ($n = 204$) *N* (%)
1. Disgust			
Can you use the toilet without feeling disgust?			
Always	186 (43.9)	98 (44.5)	88 (43.1)
Sometimes	125 (29.5)	63 (28.6)	62 (30.4)
Rarely	28 (6.6)	15 (6.8)	13 (6.4)
Never	85 (20.0)	44 (20.0)	41 (20.1)
2. Health			
Can you use the toilet without worrying that it spreads diseases?			
Always	174 (41)	96 (43.6)	78 (38.2)
Sometimes	127 (30.0)	64 (29.1)	63 (30.9)
Rarely	30 (7.1)	14 (6.4)	16 (7.8)
Never	93 (21.9)	46 (20.9)	47 (23.0)
3. Privacy			
Can you use the toilet in private, without being seen?			
Always	249 (58.7)	134 (60.9)	115 (56.4)
Sometimes	94 (22.2)	45 (20.5)	49 (24.0)
Rarely	15 (3.5)	7 (3.2)	8 (3.9)
Never	66 (15.6)	34 (15.5)	32 (15.7)
4. Shame			
Can you use the toilet without feeling ashamed for any reason?			
Always	198 (46.7)	105 (47.7)	93 (45.6)
Sometimes	113 (26.7)	60 (27.3)	53 (26.0)
Rarely	28 (6.6)	17 (7.7)	11 (5.4)
Never	85 (20.0)	38 (17.3)	47 (23.0)
5. Safety			
Are you able to feel safe while using the toilet?			
Always	193 (45.5)	90 (40.9)	103 (50.5)
Sometimes	117 (27.6)	63 (28.6)	54 (26.5)
Rarely	34 (8.0)	19 (8.6)	15 (7.4)
Never	80 (18.9)	48 (21.8)	32 (15.7)

Note: No respondents selected "prefer not to answer" for any attribute.

Abbreviation: SanQoL, sanitation-related quality of life.

We observed convergence between SanQoL index values and the WHO-5 mental well-being index, and between two respondents using the same toilet. The ICC for test-retest reliability was well above commonly used thresholds.

The benefits of sanitation beyond infectious disease are likely to underpin household WTP for sanitation improvements, but are not captured in DALYs or QALYs. The scale of the sanitation challenge, with 2 billion people lacking a basic service and US$ 20 billion invested annually in LMICs, makes it all the more surprising there been no patient/person-reported outcome measure for use in economic evaluations (Hutton & Chase, 2016). Our measure addresses this by quantifying the QoL benefits of sanitation in a way which can be applied in economic evaluation. The aggregation into a single score weighted by the relative value of attributes elicited from the target population underlies the claim to be measuring the value of sanitation overall, not only the status of diverse attributes. Our interim ranking-based approach to valuation is straightforward for respondents to understand, but it is limited in that encourages trading off of attributes as a whole rather than trading off different levels of attributes. As a consequence, the rank sum method assumes equal intervals between response categories, which does not capture relative value to a sufficient degree (Brazier et al., 2016). In

TABLE 8 Associations between SanQoL attribute scores and user/toilet characteristics

	P-values on coefficients in GLMM				
Variable	**Disgust**	**Health**	**Shame**	**Safety**	**Privacy**
User characteristics					
Female	0.698	0.461	0.250	<0.001***	0.781
Aged 60+	0.594	0.757	0.336	0.283	0.795
Wealth index	0.028**	0.441	0.855	0.198	0.816
Toilet characteristics					
Toilet has concrete or tile floor	0.271	0.003***	0.006***	<0.001***	0.003***
Toilet has masonry or complete zinc sheet walls	0.155	0.184	0.841	0.014**	0.030**
Toilet has inside lock	0.118	0.981	0.080*	0.005***	0.003***
Enumerator smells faeces on entering toilet	0.005***	<0.001***	<0.001***	0.171	0.084*

Note: *, **, *** indicate significance at the 10%, 5% and 1% level. All presence associations at the 5% level or below were in the hypothesised direction. Full regression output is in Supplementary Material G.

Abbreviations: GLMM, generalised linear mixed models; SanQoL, sanitation-related quality of life.

its place, a DCE-based approach is planned, which we were unable to apply in this study, since constraints within a PhD research budget precluded a separate valuation survey until items were finalised and validity and reliability investigated.

In combination with a measure of time using a sanitation service, the index values arising from our instrument can quantify incremental benefits in a sanitation-focused cost-effectiveness analysis (CEA). For example, analogous to the QALY, a year of sanitation service experienced at full sanitation-related capability (i.e., SanQoL = 1) might be used in CEA. The "length" dimension would be person years of service rather than life years. However, since CBA is most commonly used to evaluate sanitation interventions (Hutton & Chase, 2016), it would be important to estimate the monetary value of a such a weighted person year, based on WTP as with QALYs (Ryen & Svensson, 2015). After monetary valuation, the broader QoL benefits could be summed with health gains and other benefits such as time savings. It is through CBA that SanQoL is likely to have relevance to decision-makers with broader resource allocation remits (e.g., Ministry of Finance, Health, or Education) through comparing sanitation to non-sanitation interventions. See for example the Copenhagen Consensus (2020) studies using CBA to prioritise amongst 80+ interventions across multiple sectors in countries including Ghana, Haiti, Bangladesh, and India (Ghana Priorities, 2020). In theory, techniques for mapping to preference-based HRQoL measures could be used to estimate SanQoL-related QALYs gained from sanitation interventions, in addition to QALY gains from avoided infectious disease (Chuang & Whitehead, 2012). However, this may be conceptually problematic as the scope of SanQoL is so different to HRQoL measures, and would not be a research priority until QALYs are regularly used to evaluate sanitation interventions.

Measuring objective toilet characteristics has often been the focus of efforts to assess sanitation quality (Schelbert et al., 2020; Tidwell et al., 2018). Instead, SanQoL focuses on individuals' self-assessed capabilities. This is important because sanitation interventions might improve toilets but with heterogeneous impacts on QoL, particularly if appropriate attention is not given to gender and social norms (O'Reilly, 2016). The only previous attempt to measure aspects of these outcomes was a 60-item profile measure of women's sanitation insecurity (WSI) developed in rural India (Caruso et al., 2017). It is not appropriate for use in economic evaluation due to its lack of an overall score, use of equal weighting, and focus on women only. Its length also makes it impractical for routine use. Finally, while the WSI profile is conceptualised and applied as a measure of exposure (Caruso et al., 2018; Delea et al., 2019), SanQoL focuses on capabilities as QoL outcomes.

A short-form approach, with one item per attribute, was necessary for permitting future valuation through preference elicitation. In support of content validity, we note that our SanQoL attributes are broadly consistent with studies of insecurity, stress and motives related to sanitation in both rural and urban areas (Novotný et al., 2018; Sclar et al., 2018). However, there are exceptions. For example, convenience is commonly identified as important in settings where open defecation (OD) is prevalent (Novotný et al., 2018). However, OD is uncommon in our setting and most urban settings in general (WHO & UNICEF, 2021). In other settings where use of public toilets is common, convenience may also be an important benefit of switching to a household toilet. However, most economic evaluations of sanitation interventions already capture a major aspect of convenience by including the value of time savings (Hutton & Chase, 2016). Since the SanQoL attributes align with those identified in studies across diverse settings, the measure is likely to be appropriate for

widespread use. However, in any setting, further assessment of validity through cognitive interviews is recommended, as recently undertaken for SanQoL in urban Ghana (Tidwell et al., 2020). In settings with substantially different sanitation practices, for example, rural areas where open defecation is common, further qualitative work may be required. Respondents not using toilets may need questions reformulated in terms of "Can you practise open defecation..." rather than "Can you use the toilet...."

Across the five SanQoL items, only 6% of responses (range: 4%–8%) endorsed the "rarely" option (Table 7). This compares to 11% in piloting. Alternative response options could be investigated in future research in a broader population, to establish the effect of response labels on endorsement distributions. However, our distributions align with the fact that for the WSI profile, also on a four-level frequency scale, only 5% of respondents endorsed its "often" category in a study with three times the sample size (Caruso et al., 2017). It may be that respondents prefer to scale these concepts on three levels. However, we think it is more likely that "sometimes" (used in both SanQoL and the WSI profile) comprises a very broad category, such that it might dominate the other middle category in any four-level scale. This could be investigated using think-aloud methods or item response theory in a larger sample of a population using a broad range of toilet types.

Use of the intervention and control groups of a trial as the study population was pragmatic, and contributed to the objectives of our broader body of research. However, it did provide only a narrow set of sanitation service types and living arrangements–for example, almost all respondents shared their toilet with other households in a compound setting. In addition, our study's sample size was relatively small. Further research on the validity and measurement properties of SanQoL should take place in other settings, and a larger random sample of a target population that uses a broader range of sanitation service levels. In addition, our sample provides evidence of validity only for those aged 18 years or older, and further research would be required for use in younger populations who may experience sanitation in different ways (Nallari, 2015; Sahoo et al., 2015).

5 | CONCLUSION

The SanQoL instrument provides a measure of sanitation-related quality of life, and we have provided evidence for its validity and reliability in this setting. It enables measurement and valuation of the attributes of sanitation which users consider important, weighted by the relative value they place on those attributes. Such a measure can support future cost-effectiveness and cost-benefit analyses, as well as being appropriate for impact evaluation or routine program monitoring to assess the performance of interventions. The results of validity and reliability investigations reported here provide evidence that this measure can be used for assessing SanQoL in urban settings of LMICs, but further testing in other settings is required.

ACKNOWLEDGMENTS

We greatly appreciated the cooperation of survey respondents. We also appreciated the efforts of the fieldworkers: Euclimia Titosse, Carla Panguene, João-Pedro Guambe and Faustino Benzane. The authors gratefully acknowledge the technical and logistical support received from Vasco Parente and his team at WSUP Mozambique. We also benefitted from discussions with Sarah Smith, Julia Lohmann, Catherine Pitt and seminar participants at the Global Health Economics Centre at LSHTM. The inputs of the expert reviewers were helpful in improving items. This work was supported by the Economic and Social Research Council through a PhD studentship. The fieldwork was funded by the Bill & Melinda Gates Foundation. The funders had no role in the identification, design, conduct, or reporting of the analysis.

CONFLICT OF INTERESTS

The time or research costs of the following researchers on this study was funded in part by the study's funders, the Bill and Melinda Gates Foundation: Ian Ross, Oliver Cumming, Zaida Adriano, Rassul Nala, Joe Brown, Robert Dreibelbis.

DATA AVAILABILITY STATEMENT

Deidentified individual participant data, codebook, and replication code are available on LSHTM data repository: https://doi.org/10.17037/DATA.00002442.

ORCID

Ian Ross https://orcid.org/0000-0002-2218-5400

REFERENCES

Baker, R., Mason, H., McHugh, N., & Donaldson, C. (2021). Public values and plurality in health priority setting: What to do when people disagree and why we should care about reasons as well as choices. *Social Science & Medicine*, *277*(April), 113892. https://doi.org/10.1016/j.socscimed.2021.113892

Boardman, A., Greenberg, D., Vining, A., & Weimer, D. (2018). *Cost-benefit analysis: Concepts and practices* (5th ed.). Cambridge University Press.

Bowden, A., Fox-Rushby, J. A., Nyandieka, L., & Wanjau, J. (2002). Methods for pre-testing and piloting survey questions: Illustrations from the KENQOL survey of health-related quality of life. *Health Policy and Planning*, *17*(3), 322–330. https://doi.org/10.1093/heapol/17.3.322

Brazier, J., Ratcliffe, J., Salomon, J., & Tsuchiya, A. (2016). *Measuring and valuing health benefits for economic evaluation*. Oxford University Press.

Brown, J., Cumming, O., Bartram, J., Cairncross, S., Ensink, J., Holcomb, D., & Schmidt, W. P. (2015). A controlled, before-and-after trial of an urban sanitation intervention to reduce enteric infections in children: Research protocol for the Maputo Sanitation (MapSan) study, Mozambique. *BMJ Open*, *5*(6), 1–12.

Caruso, B. A., Clasen, T., Yount, K. M., Cooper, H. L. F., Hadley, C., & Haardörfer, R. (2017). Assessing women's negative sanitation experiences and concerns: The development of a novel sanitation insecurity measure. *International Journal of Environmental Research and Public Health*, *14*(7), 1–22. https://doi.org/10.3390/ijerph14070755

Caruso, B. A., Cooper, H. L. F., Haardörfer, R., Yount, K. M., Routray, P., Torondel, B., & Clasen, T. (2018). The association between women's sanitation experiences and mental health: A cross-sectional study in rural, Odisha India. *SSM – Population Health*, *5*(March), 257–266. https://doi.org/10.1016/j.ssmph.2018.06.005

Castel-Branco, R., & Andrés, R. V. (2019). *Towards universal social security for the elderly in Mozambique*. International Labour Organisation.

Chuang, L. H., & Whitehead, S. J. (2012). Mapping for economic evaluation. *British Medical Bulletin*, *101*(1), 1–15. https://doi.org/10.1093/bmb/ldr049

Cicchetti, D. V. (1994). *Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment*, 6, 284–290. American Psychological Association. https://doi.org/10.1037/1040-3590.6.4.284

CMM. (2017). *Postura de Saneamento e Drenagem, Pub. L. No. 68/AMM/2016*. Assembleia Municipal de Maputo.

Coast, J. (2019). Assessing capability in economic evaluation: A life course approach? *The European Journal of Health Economics*, *20*(6), 779–784. https://doi.org/10.1007/s10198-018-1027-6

Coast, J., Flynn, T. N., Natarajan, L., Sproston, K., Lewis, J., Louviere, J. J., & Peters, T. J. (2008). Valuing the ICECAP capability index for older people. *Social Science & Medicine*, *67*(5), 874–882. https://doi.org/10.1016/j.socscimed.2008.05.015

Coast, J., Kinghorn, P., & Mitchell, P. (2015). The development of capability measures in health economics: Opportunities, challenges and progress. *Patient*, *8*(2), 119–126. https://doi.org/10.1007/s40271-014-0080-1

Coast, J., Smith, R. D., & Lorgelly, P. (2008). Welfarism, extra-welfarism and capability: The spread of ideas in health economics. *Social Science & Medicine*, *67*(7), 1190–1198. https://doi.org/10.1016/j.socscimed.2008.06.027

Cookson, R. (2003). Willingness to pay methods in health care: A sceptical view. *Health Economics*, *894*, 891–894. https://doi.org/10.1002/hec.847

Copenhagen Consensus. (2020). Copenhagen Consensus projects. Retrieved September 4, 2020, from https://www.copenhagenconsensus.com/projects

de Kruijk, H., & Rutten, M. (2007). *Weighting dimensions of poverty based on peoples priorities: Constructing a Composite Poverty Index for the Maldives*. Institute for International & Development Economics.

Delea, M. G., Snyder, J. S., Belew, M., Caruso, B. A., Garn, J. V., Sclar, G. D., & Freeman, M. C. (2019). Design of a parallel cluster-randomized trial assessing the impact of a demand-side sanitation and hygiene intervention on sustained behavior change and mental well-being in rural and peri-urban Amhara, Ethiopia: Andilaye study protocol. *BMC Public Health*, *19*(1), 1–15. https://doi.org/10.1186/s12889-019-7040-6

Drummond, M., Stoddard, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes (fourth)*. Oxford University Press.

Elmendorf, M., & Buckles, P. K. (1980). Sociocultural aspects of water supply and excreta disposal (volume 5). In *Appropriate technology for water supply and sanitation*. World Bank.

Euroqol Group. (2009). EuroQol group EQ-5D™ Health Questionnaire. *3*. http://Www.Euroqol.Org/Eq-5D-Products/How-To-Obtain-Eq-5D.Html

Fayers, P., & Machin, D. (2015). *Quality of life: The assessment, analysis and interpretation of patient-reported outcomes* (3rd ed.). Wiley.

Ghana Priorities (2020). *Ghana priorities eminent panel findings*.

Goodwin, E., & Green, C. (2016). A systematic review of the literature on the development of condition-specific preference-based measures of health. *Applied Health Economics and Health Policy*, *14*(2), 161–183. https://doi.org/10.1007/s40258-015-0219-9

Greco, G. (2016). Setting the weights: The women's capabilities index for Malawi. *Social Indicators Research*, *135*(2), 457–478. https://doi.org/10.1007/s11205-016-1502-3

Greco, G., Lorgelly, P., & Yamabhai, I. (2016). Outcomes in economic evaluations of public health interventions in low- and middle-income countries: Health, capabilities and subjective wellbeing. *Health Economics*, *25*(1), 83–94. https://doi.org/10.1002/hec.3302

Groce, N., Bailey, N., Lang, R., Trani, J. F., & Kett, M. (2011). Water and sanitation issues for persons with disabilities in low- and middle-income countries: A literature review and discussion of implications for global health and international development. *Journal of Water and Health*, *9*(4), 617–627. https://doi.org/10.2166/wh.2011.198

Gruebner, O., Khan, M. M. H., Lautenbach, S., Müller, D., Krämer, A., Lakes, T., & Hostert, P. (2012). Mental health in the slums of Dhaka - a geoepidemiological study. *BMC Public Health*, *12*(1), 177. https://doi.org/10.1186/1471-2458-12-177

Hawkins, P., & Muximpua, O. (2015). *Developing business models for fecal sludge management in Maputo*. World Bank.

Hensher, D., Rose, J., & Greene, W. (2015). *Applied choice analysis* (2nd ed.). Cambridge University Press.

Hutton, G., & Chase, C. (2016). The knowledge base for achieving the sustainable development goal targets on water supply, sanitation and hygiene. *International Journal of Environmental Research and Public Health*, *13*(6), 536. https://doi.org/10.3390/ijerph13060536

INE. (2019). *IV Recenseamento Geral da População e Habitação, 2017: Resultados Definitivos*. Instituto Nacional de Estatística.

Jenkins, M., & Curtis, V. (2005). Achieving the "good life": Why some people want latrines in rural Benin. *Social Science & Medicine*, *61*(11), 2446–2459. https://doi.org/10.1016/j.socscimed.2005.04.036

Karimi, M., & Brazier, J. (2016). Health, health-related quality of life, and quality of life: What is the difference? *PharmacoEconomics*, *34*(7), 645–649. https://doi.org/10.1007/s40273-016-0389-9

Knee, J., Sumner, T., Adriano, Z., Anderson, C., Bush, F., Capone, D., Casmo, V., Holcomb D., Kolsky P., MacDougall A., Molotkova E., Braga J. M., Russo C., Schmidt W. P., Stewart J., Zambrana W., Zuin V., Nalá R., Cumming O., & Brown J. (2021). Effects of an urban sanitation intervention on childhood enteric infection and diarrhoea in Mozambique. *ELife*, *10*. e62278. https://doi.org/10.7554/eLife.62278

Konerding, U. (2013). What does Cronbach's alpha tell us about the EQ-5D? A methodological commentary to "psychometric properties of the EuroQol Five-Dimensional Questionnaire (EQ-5D-3L) in caregivers of autistic children. *Quality of Life Research*, *22*(10), 2939–2940. https://doi.org/10.1007/s11136-013-0430-9

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Mangham, L. J., Hanson, K., & McPake, B. (2009). How to do (or not to do)...Designing a discrete choice experiment for application in a low-income country. *Health Policy and Planning*, *24*(2), 151–158. https://doi.org/10.1093/heapol/czn047

MOPH. (2011). *Estratégia Nacional de Água e Saneamento Urbano*.

Mukherjee, N. (2001). *Achieving sustained sanitation for the poor: Policy and strategy lessons from participatory assessments in Cambodia, Indonesia and Vietnam*. Retrieved from World Bank.

mWater. (2019). *mWater Surveyor*.

Nallari, A. (2015). "All we want are toilets inside our homes!": The critical role of sanitation in the lives of urban poor adolescent girls in Bengaluru, India. *Environment and Urbanization*, *27*(1), 73–88. https://doi.org/10.1177/0956247814563514

Novotný, J., Hasman, J., & Lepič, M. (2018). Contextual factors and motivations affecting rural community sanitation in low- and middle-income countries: A systematic review. *International Journal of Hygiene and Environmental Health*, *221*(2), 121–133. https://doi.org/10.1016/j.ijheh.2017.10.018

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw Hill.

O'Reilly, K. (2016). From toilet insecurity to toilet security: Creating safe sanitation for women and girls. *Wiley Interdisciplinary Reviews: Water*, *3*(1), 19–24. https://doi.org/10.1002/wat2.1122

Peasgood, T., Brazier, J., Mukuria, C., & Rowen, D. (2014). *A conceptual Comparison of well-being measures Used in the UK (No. 026)*. Policy Research Unit in Economic Evaluation of Health & Care Interventions (EEPRU). Retrieved from http://www.eepru.org.uk/wp-content/uploads/2017/11/eepru-report-a-conceptual-comparison-of-well-being-measures-sept-2014-026.pdf

Prüss-Ustün, A., Wolf, J., Bartram, J., Clasen, T., Cumming, O., Freeman, M. C., Gordon, B., Hunter, P. R., Medlicott, K., & Johnston, R. (2019). Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: An updated analysis with a focus on low- and middle-income countries. *International Journal of Hygiene and Environmental Health*, *222*(5), 765–777. https://doi.org/10.1016/j.ijheh.2019.05.004

Reeve, B. B., Wyrwich, K. W., Wu, A. W., Velikova, G., Terwee, C. B., Snyder, C. F., Schwartz, C., Revicki, D. A., Moinpour, C. M., McLeod, L. D., Lyons, J. C., Lenderking, W. R., Hinds, P. S., Hays, R. D., Greenhalgh, J., Gershon, R., Feeny, D., Fayers, P. M., Cella, D., & Butt, Z. (2013). ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of Life Research*, *22*(8), 1889–1905. https://doi.org/10.1007/s11136-012-0344-y

Ross, I. (2021). *Data for: "Sanitation-related quality of life (SanQoL) in Maputo, Mozambique"*. LSHTM Data Compass. https://doi.org/10.17037/DATA.00002442

Ross, I., Cumming, O., Dreibelbis, R., Adriano, Z., Nala, R., & Greco, G. (2021). How does sanitation influence people's quality of life? Qualitative research in low-income areas of Maputo, Mozambique. *Social Science & Medicine*, *272*, 113709. https://doi.org/10.1016/j.socscimed.2021.113709

Ryen, L., & Svensson, M. (2015). The willingness to pay for a quality adjusted life year: A review of the empirical literature. *Health Economics*, *24*(10), 1289–1301. https://doi.org/10.1002/hec.3085

Sahoo, K. C., Hulland, K., Caruso, B. A., Swain, R., Freeman, M. C., Panigrahi, P., & Dreibelbis, R. (2015). Sanitation-related psychosocial stress: A grounded theory study of women across the life-course in Odisha, India. *Social Science & Medicine*, *139*, 80–89. https://doi.org/10.1016/j.socscimed.2015.06.031

Schelbert, V., Meili, D., Alam, M.-U., Simiyu, S., Antwi-Agyei, P., Adjei, K. A., Dwumfour-Asare, B., Rahman, M., Ferdous, S., Sarker, S., Günther, I., & Lüthi, C. (2020). When is shared sanitation acceptable in low-income urban settlements? A user perspective on shared sanitation quality in Kumasi, Kisumu and Dhaka. *Journal of Water, Sanitation and Hygiene for Development*, *10*, 1–968. https://doi.org/10.2166/washdev.2020.084

Sclar, G. D., Penakalapati, G., Caruso, B. A., Rehfuess, E. A., Garn, J. V., Alexander, K. T., Freeman, M. C., Boisson, S., Medlicott, K., & Clasen, T. (2018). Exploring the relationship between sanitation and mental and social well-being: A systematic review and qualitative synthesis. *Social Science & Medicine*, *217*, 121–134. https://doi.org/10.1016/j.socscimed.2018.09.016

Sen, A. (1980). Equality of what? In S. M. McMurrin (Ed.), *The Tanner lecture on human values* (vol. 1, pp. 197–220). Cambridge University Press. https://doi.org/10.1093/0198289286.003.0002

Sen, A. (1985). *Commodities and capabilities*. Elsevier Science Publishers.

Sen, A. (1993). Capability and well-being. In A. Sen & M. Nussbaum (Eds.), *The quality of life* (pp. 30–53). Clarendon press.

Simon, J., Anand, P., Gray, A., Rugkåsa, J., Yeeles, K., & Burns, T. (2013). Operationalising the capability approach for outcome measurement in mental health research. *Social Science & Medicine*, *98*, 187–196. https://doi.org/10.1016/j.socscimed.2013.09.019

StataCorp. (2019). *Stata statistical software: Release 16*. StataCorp LLC.

Stillwell, W. G., Seaver, D. A., & Edwards, W. (1981). A comparison of weight approximation techniques in multiattribute utility decision making. *Organizational Behavior & Human Performance*, *28*(1), 62–77. https://doi.org/10.1016/0030-5073(81)90015-5

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford University Press.

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Tidwell, J. B., Chipungu, J., Chilengi, R., & Aunger, R. (2018). Assessing peri-urban sanitation quality using a theoretically derived composite measure in Lusaka, Zambia. *Journal of Water, Sanitation and Hygiene for Development*, *8*(October), 1–678. https://doi.org/10.2166/washdev.2018.029

Tidwell, J. B., Nyarko, K. B., Ross, I., Dwumfour-Asare, B., & Scott, P. (2020). Evaluation of user experiences for the clean team Ghana container-based sanitation service in Kumasi, Ghana. *MedRxiv*. https://doi.org/10.1101/2020.10.23.20218578

Tilley, E., Bieri, S., & Kohler, P. (2013). Sanitation in developing countries: A review through a gender lens. *Journal of Water, Sanitation and Hygiene for Development*, *3*(3), 298–314. https://doi.org/10.2166/washdev.2013.090

Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 well-being index: A systematic review of the literature. *Psychotherapy and Psychosomatics*, *84*(3), 167–176. https://doi.org/10.1159/000376585

UN-HABITAT. (2010). *Mozambique cities profile*. Retrieved from http://mirror.unhabitat.org/%0Apmss/getElectronicVersion.aspx?nr=3002&alt=1

Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning*, *21*(6), 459–468. https://doi.org/10.1093/heapol/czl029

Ware, J., Brook, R., Williams, K., Stewart, A., & Davies-Avery, A. (1980). *Conceptualisation and measurement of health for adults in the health insurance study. Vol. 1: Model of health and methodology*. Rand Corporation.

Weller, S. C., & Romney, A. K. (1988). *Systematic data collection*. Sage Qualitative Research Methods Series #10.

WHO & UNICEF. (2021). *Progress on household drinking water, sanitation and hygiene 2000–2020 five years into the SDGs*. Retrieved from https://washdata.org/sites/default/files/2021-06/jmp-2021-wash-households-LAUNCH-VERSION.pdf

WHO. (1948). *Constitution of the World health Organization* (ISBN: 12571729).

WHO. (2017). *UN-water Global Analysis and Assessment Of Sanitation and drinking-water (GLAAS) 2017*. CC BY-NC-SA 3.0 IGO.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Ross, I., Greco, G., Opondo, C., Adriano, Z., Nala, R., Brown, J., Dreibelbis, R., & Cumming, O. (2021). Measuring and valuing broader impacts in public health: Development of a sanitation-related quality of life instrument in Maputo, Mozambique. *Health Economics*, 1–15. https://doi.org/10.1002/hec.4462