Measuring and valuing quality of life in the economic evaluation of sanitation interventions

Ian Ross

Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy of the University of London

March 2021

Department of Global Health and Development
Faculty of Public Health and Policy
London School of Hygiene and Tropical Medicine

Funded by the Economic and Social Research Council, and supported by grants from the Bill & Melinda Gates Foundation

Research group affiliations: Centre for Health Economics in London; Environmental Health Group.

Declaration

I, Ian Ross, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Ian Ross

March 2021

Abstract

Recent trials of basic sanitation interventions in low- and middle-income countries have identified little or no health impact, despite improving access to toilets and their quality. However, qualitative studies frequently report that toilet users value broader benefits for privacy, safety and status. Economic evaluations have not included these benefits, in the absence of means to measure them quantitatively, potentially leading to misallocated resources. The aim of this research was to develop and apply an approach to measuring and valuing quality of life in the economic evaluation of sanitation interventions. This is a paper-style thesis which incorporates five papers linked by short pieces of supporting material.

By integrating qualitative and quantitative methods from health economics, this thesis outlines the development and application of a measure of "sanitation-related quality of life" (SanQoL). The thesis finds that the benefits of an urban sanitation intervention for toilet users' quality of life can be quantitatively measured and valued, working alongside the Maputo Sanitation trial in Mozambique. Attributes of the measure are first identified in qualitative research using the capability approach. The validity and reliability of the ensuing SanQoL measure are assessed using psychometric analytic methods.

The SanQoL measure captures the degree of achievement of five sanitation-related capabilities: privacy, safety, health, shame and disgust. Rescaling with user-derived weights results in SanQoL index values ranging from zero (no sanitation capability) to one (full sanitation capability). These index values can be used to weight sanitation-adjusted person years (SAPYs), a proposed measure of the value of sanitation. After estimating the effect of a shared, urban sanitation intervention on SanQoL, that effect is applied in a cost-effectiveness analysis using the novel SAPY measure as the outcome. This thesis demonstrates how measuring and valuing toilet users' quality of life brings new insights to the effectiveness and cost-effectiveness of sanitation interventions.

Acknowledgments

I would not have been unable to undertake this research without the support of my wife Lucy and “diversion” offered by my two young boys, the older of whom was born the week I started this PhD. They helped me in keeping regular, reasonable hours and in retaining a focus on what I have reason to value.

Giulia Greco and Oliver Cumming have been excellent supervisors, always providing sound advice. They helped me bridge the worlds of health economics and WASH, and the worlds of “non-academia” and academia. My writing is substantially better after relentless flagging of imprecision and unreferenced assertions. I have also appreciated the sage advice of Graham Medley at a strategic level. My advisory team of Catherine Pitt, Charles Opondo and Robert Dreibelbis have been pillars of good sense in their respective fields of expertise, when I came to them with assorted methodological problems. Being based in a village outside Oxford has meant I’ve only made it to LSHTM once a fortnight at best. I am grateful therefore to the members of the environmental health group (especially Team 411) and the economic evaluation group in the Centre for Health Economics in London, for making me feel extremely welcome when I did manage to join them at base camp. Both groups were important sounding boards for my ideas throughout the research, and their regular meetings were a source of learning and inspiration.

Thank you to colleagues in Mozambique, particularly Zaida Adriano, for dedicated collaboration (and putting up with my intermediate Portuguese). The support of my Portuguese teacher Angela Arsenio in fortnightly skype lessons throughout my study was crucial. I was fortunate to work with almost the same fieldworker team for both qualitative and quantitative studies. Their enthusiasm and flexibility was invaluable – thank you to Euclimia Titosse, Neusia Chavate, João-Pedro Guambe, Gimo Jenga, Carla Panguene and Faustino Benzane. Thanks also to the team at WSUP Mozambique, particularly Vasco Parente, for answering my many questions about their sanitation intervention.

This thesis has benefited from the constructive critique of participants at the Health Economics Study Group meeting in Newcastle (2020), the UNC Water and Health conference (2019) and Human Development and Capabilities Association conference (2019). Finally, the financial support of an Economic and Social Research Council +3 studentship was a necessary condition for embarking on this research. The funding for fieldwork costs from the Bill and Melinda Gates Foundation is also gratefully acknowledged.

In memoriam

This thesis is dedicated to the memory of Prof. Val Curtis, Director of the Environmental Health Group at LSHTM, who died a few weeks before I submitted this thesis. Her work on motives underlying behaviour is foundational in any investigation of what people have reason to value about sanitation.

Table of Contents

List of tables

Chapter 8

Chapter 9

List of figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Abbreviations

BCR	Benefit-cost ratio
CBA	Cost-benefit analysis
CEA	Cost-effectiveness analysis
CMM	*Conselho Municipal de Maputo*
CSB	Community sanitation block
DALY	Disability-adjusted life year
DSA	Deterministic sensitivity analysis
EQ-5D	EuroQoL five-dimension measure
FGD	Focus group discussion
HRQoL	Health-related quality of life
HWT	Household water treatment
ICER	Incremental cost-effectiveness ratio
IDI	In-depth interview
LMIC	Low- and middle-income countries
MZN	Mozambican meticais
OD	Open defecation
PL	Pit latrine
PSA	Probabilistic sensitivity analysis
QALY	Quality-adjusted life year
QoL	Quality of life
SAPY	Sanitation-adjusted person year
SanQoL	Sanitation-related quality of life
ST	Shared toilet
VAS	Visual analogue scale
WASH	Water, sanitation and hygiene
WHO-5	WHO mental wellbeing scale in five dimensions
WSUP	Water and Sanitation for the Urban Poor

Chapter 1: Introduction

In this chapter, I first motivate the thesis briefly, by explaining the rationale for economic evaluation in general, introducing how economic evaluation has been used in the sanitation sector, and highlighting the main gaps in outcome measurement and valuation. These topics are discussed in more depth in the literature review (Chapter 2). Next, I introduce the research setting in Maputo, Mozambique, and explain the research aim and objectives. Finally, I outline the structure of the thesis, and summarise ethical approval and funding.

1.1. Motivation for the thesis

Many basic services such as health care, education and sanitation can have positive externalities (Pigou, 1920). Everyone benefits when their neighbours are immunised, educated, and using a toilet, forming the justification for the use of public funds. Considering sanitation in particular, an excreta-free environment is a public good, in that it is non-rival and non-excludable (Samuelson, 1954). An excreta-free environment is more likely when everyone uses toilets and their faecal waste is safely contained and treated. While there might be a free market for toilet components, the characteristics of sanitation services are such that there is not a free market for an excreta-free environment (Batley and Mcloughlin, 2015). Together, these considerations form the justification for the use of public funds for sanitation investments in particular.

Since public funds are a scarce resource, particularly in low- and middle-income countries (LMICs), decisions must be made about their allocation. Economic evaluations aim to inform the efficient allocation of resources, by comparing the costs and outcomes of interventions (Drummond et al., 2015). In most sectors, the dominant approach to economic evaluation is "welfarist" cost-benefit analysis (CBA), undertaking monetary valuation of benefits. Such an approach is common in project appraisal by development banks (World Bank, 2010). The health sector is unusual in its preference for an "extra-welfarist" approach, which focuses on maximising one source of value only, health in this case (Culyer, 1989). Cost-effectiveness analysis (CEA) is therefore the method of choice in health, with benefits captured by some measure of the value of health rather than in monetary terms. In high-income countries, the value of health is usually measured by the quality-adjusted life year (QALY), while in LMICs the disability-adjusted life year (DALY) is more commonly used (Drummond et al., 2015).

In the economic evaluation of sanitation, both welfarist (CBA) and extra-welfarist (CEA) approaches have been applied. The earliest forays into economic evaluation of sanitation interventions were by health professionals (Cvjetanovic and Grab, 1976), using CEA with averted cases and deaths as the outcome (Walsh and Warren, 1979). However, when compared to health interventions on a health basis only, interventions such as sanitation and water supply seemed like a “bad” investment (Briscoe, 1984). This is because a large number of impacts are outside health, for example in the value of avoided travel time to collect water or practice open defecation (Feachem, 1986). This has meant that CBA has been the more popular method for economic evaluation of sanitation programmes (Hutton and Chase, 2016; Whittington et al., 2009).

Sanitation CBA studies have most commonly measured and valued travel time savings and health benefits (Hutton and Chase, 2016). They have routinely excluded some likely benefits, in the absence of methods for their measurement and valuation (Hutton et al., 2020). These include: (i) improvements in quality of life in the use of toilets, beyond disease reduction, such as privacy, safety and dignity, (ii) welfare gains beyond the use of toilets, such as an excreta-free environment being more pleasant to live in, (iii) positive externalities from safe containment, conveyance and treatment of excreta. These exclusions may result in misallocation of resources if interventions differ in the extent to which they produce these benefits, as appears likely. The exclusion of the quality of life outcomes is particularly problematic as these are often cited by toilet users as the most important perceived benefits or motives for investment (Jenkins and Curtis, 2005; Jenkins and Scott, 2007; Mukherjee, 2001). As such, they are likely to underpin household willingness to pay for sanitation improvements, and thus be important in welfarist CBA. To date however, benefits for quality of life have been seen as “intangible” (Hutton and Chase, 2016).

In this thesis, I develop methods to measure and value these quality of life (QoL) outcomes such as privacy, safety and dignity. While there have been willingness to pay studies for toilet technologies or other sanitation services, I am not aware of any attempts to value QoL outcomes in CBA, which I confirm in a systematic review reported in Chapter 3. In this thesis, I develop a means for measuring sanitation-related quality of life, grounded in qualitative research. I then apply it in a cost-effectiveness analysis (CEA) by proposing an extra-welfarist measure of the value of sanitation: the sanitation-adjusted person year (SAPY). CEA using SAPYs can be used to compare sanitation interventions. For comparisons beyond sanitation, the monetary value of a SAPY can be elicited and applied in welfarist CBA.

1.2. Research aim, objectives and overall design

Aim

The aim of this research was to develop and test an approach to economic evaluation of sanitation interventions by measuring and valuing quality of life.

Objectives

In order to achieve this aim, this research sought to achieve the following objectives:

1. To assess the extent to which quality of life outcomes have been measured and valued in economic evaluations of sanitation and drinking water interventions.
2. To develop a psychometric measure of quality of life related to sanitation, based on qualitative research.
3. To estimate the effect of access to better toilets on quality of life, in a quantitative study using the measure.
4. To estimate the cost-effectiveness of a shared sanitation intervention by valuing quality of life gains.

Overall research design and methods

To achieve these objectives, I designed a body of mixed methods research. While the primary disciplinary perspective was health economics, I also drew on methods from psychology, anthropology and epidemiology. Throughout, the design was informed by the capability approach, which has its roots in philosophy as well as economics. Starting from a systematic review and qualitative methods, the study transitioned towards quantitative methods through the course of the research, as follows:

Systematic review: I used systematic review methods to synthesise evidence on the cost-benefit and cost-effectiveness of drinking water and sanitation interventions in LMICs. This drew on best practice for systematic reviews of economic evaluations.

Qualitative research: I used in-depth interviews and focus group discussions to identify what people most valued about sanitation in this setting, drawing together findings using framework analysis. In order to triangulate findings on attributes' relative importance, I used pile-sorting and triad methods from cognitive anthropology.

Measure development: The design of the descriptive system of the measure drew on approaches to measurement and valuation of quality of life and capabilities from health economics. I refined the quantitative survey instrument using piloting and cognitive interviews, to explore whether and how people had understood the questions. Marking the transition to quantitative methods, the assessment of different aspects of validity and reliability required psychometric analytical approaches from psychology.

Effect estimation: I combined empirical strategies from econometrics and epidemiology to evaluate the effect of using higher-quality toilets on quality of life, using generalised linear mixed models to account for clustering.

Costing and cost-effectiveness: I combined bottom-up and top-down economic costing methods to estimate the incremental cost of a pour-flush toilet intervention as compared to existing use of pit latrines. The measure of the value of sanitation was based on the concept of the quality-adjusted life year from health economics. Likewise, the approach to cost-effectiveness modelling and sensitivity analysis were rooted in health economics.

1.3. Structure of the thesis

This thesis comprises a two-part literature review incorporating a systematic review "paper-style" chapter, an overview of the study setting, and four empirical paper-style chapters. It concludes with a summative discussion of the findings from all components. The content of the chapters is outlined below.

Chapters 2 and 4 comprise the literature review. Chapter 2 presents an overview of the theoretical underpinnings of economic evaluation in general and in health, as well as an overview of the capability approach. It also provides background on sanitation interventions, their economic benefits, and how these have been measured and valued in economic evaluations.

Chapter 3 provides background on the setting, specifically the Mozambique and Maputo broad context, the urban sanitation policy environment, the Maputo Sanitation (MapSan) Trial, and the intervention it evaluates.

Chapter 4 is a paper-style chapter comprising the second part of the literature review. It reports a systematic review of the cost-benefit and cost-effectiveness of drinking water and sanitation interventions in low- and middle-income countries. It addresses three questions: (i) what methods have economic evaluations used and with what level of quality? (ii) what conclusions can be drawn about the economic performance of different types of interventions? (iii) what types of outcomes have been measured and valued? The answer to the third question motivated the focus of this thesis on measuring sanitation-related quality of life. The rationale for including drinking water interventions was that these present similar benefit measurement and valuation challenges to sanitation, and are sometimes delivered and evaluated in combination with sanitation interventions.

Chapter 5 is a paper-style chapter which aims to investigate what people most value about sanitation in this setting, to inform a definition and conceptual model of sanitation-related quality of life. Based on in-depth interviews and focus groups using the capability approach, it identifies five core attributes of sanitation-related quality of life, or sanitation-related capabilities: avoiding disgust, health, avoiding shame, safety and privacy. These are set within a conceptual model which illustrates how sanitation interventions might improve quality of life via changes in these attributes, and how changes are likely to be mediated by individual and environmental conversion factors.

Chapter 6 is a paper-style chapter which aims to develop and assess a measure of sanitation-related quality of life (SanQoL) in the study setting. Building on the conceptual model from the qualitative research, it develops a five-item SanQoL descriptive system which is assessed using pilot and cognitive interviews. A quantitative survey (n=424) was undertaken with people living on intervention and control compounds enrolled in the MapSan trial, with data used to assess the validity and reliability of the measure. This chapter also estimates attribute weights based on a ranking exercise undertaken as part of the survey.

Chapter 7 is a paper-style chapter which aims to estimate the effect of using better-quality toilets on quality of life in the study setting. Based on the dataset collected in chapter 6, it uses generalised linear mixed models to compare SanQoL index values between people using high-quality pour-flush toilets and low-quality pit latrines. It also explores respondents' direct evaluation of their level of sanitation using a visual analogue scale, and effects on mental wellbeing as measured by the WHO-5 index.

Chapter 8 proposes a novel extra-welfarist measure of the valuation of sanitation: the sanitation-adjusted person year (SAPY). It sets out the rationale for the SAPY, its theoretical properties, its potential and its limitations.

Chapter 9 is a paper-style chapter which aims to demonstrate a novel approach to cost-effectiveness analysis of sanitation interventions, by valuing quality of life effects using SAPYs. It estimates the incremental costs and outcomes of shared toilets and community sanitation blocks delivered by an NGO intervention, as compared to existing use of shared pit latrines. Outcomes are measured in SAPYs, and compared using effectiveness estimates from Chapter 7. The incremental cost of achieving SAPY gains is calculated, and uncertainty characterised using probabilistic and deterministic sensitivity analysis.

Chapter 10 summarises the main findings across the thesis, reflects on the contribution of the research, and explores avenues for future research.

1.4. Intellectual Ownership, funding and ethical approval

This research was supported by the UK Economic and Social Research Council (ESRC) with a three-year doctoral studentship. The Bill and Melinda Gates Foundation (BMGF) provided funding for the fieldwork as part of a grant for the Maputo Sanitation Trial.

I led all elements of research in this thesis, with support and advice from my supervisors, advisory committee and upgrading examiners.

Ethical approval was granted by the Research Ethics Committee of the London School of Hygiene and Tropical Medicine (Ref: 14609) on 5th June 2018 and by the *Comité Nacional de Bioética para a Saúde* (IRB00002657) at the Ministry of Health in Mozambique on 6th November 2018. All participants in focus groups, in-depth interviews, and quantitative interviews provided written informed consent to participate. Participants were informed of their right to end discussions at any time. All audio recordings made as part of the qualitative research were permanently deleted after verification of transcripts.

Chapter 2: Literature review

2.1. Introduction

The literature review component of this thesis is in two parts. In this chapter, I address the theoretical literature which provides the context to this thesis and explain key concepts. After a description of the Mozambique setting in Chapter 3, Chapter 4 then reports a formal systematic review of cost-benefit and cost-effectiveness studies of drinking water and sanitation interventions in LMICs, presented as a paper-style chapter.

This chapter is set out in four sections. In section 2.2, I provide an overview of economic evaluation as applied to health. This includes discussing welfarism and extra-welfarism, the different types of health-adjusted life years, and comparing cost-effectiveness and cost-benefit analysis. This section of the review is mainly descriptive.

In section 2.3, I discuss economic evaluation as applied to sanitation. This starts with an explanation of levels of sanitation service, then a discussion of sanitation as a "state of being". Next, I outline types of sanitation interventions, and how the benefits of these have been measured and valued in economic evaluation. The empirical literature on the economic evaluation of sanitation is discussed as part of the systematic review in Chapter 4.

In section 2.4., I introduce theory underlying the capability approach, both in general and for health. I then discuss how the capability approach has been applied in health economic evaluation to date. Finally, I review how the capability approach has been applied in the sanitation sector. Section 2.5 briefly concludes.

2.2. Economic evaluation as applied to health

2.2.1. Welfarism and extra-welfarism as applied to health

Welfare economics is the study of the social preferability of different states of the world. At present, the neoclassical framework for welfare economics is dominant, which Hurley (2000) identifies as having four tenets (Table 1): (1) utility maximisation; (2) individual sovereignty, (3) consequentialism, and (4) welfarism.

Table 1: Four tenets of the 'welfarist' economic framework, based on Hurley (2000)

Tenet	Assumption	Nature of assumption
1. Utility maximisation	Individuals rank options by utility and "rationally" choose the highest-ranked option	Behavioural
2. Individual sovereignty	Individuals are the best judges of what contributes to their utility	Normative
3. Consequentialism	All that matters is outcomes, not the process which delivers them	Normative
4. Welfarism	Utility is all that matters in judging social welfare	Normative

Extra-welfarism loosens these restrictions, in particular the fourth, whereby the evaluative space can be widened beyond utility. However, extra-welfarism does not reject the use of utility *per se*. For example, QALYs are usually valued using preference-based measures, which themselves comprise indices derived from stated preference studies which rely on expected utility theory (von Neumann and Morgenstern, 1944).

Brouwer et al. (2008) identify four ways in which extra-welfarism is different from welfarism (Table 2). In health economic evaluation using QALYs, this is manifested in three important ways. Following the lettering in Table 2: (A) health *per se* is the maximand rather than welfare/utility (in contrast to #1 in Table 1); (B) society is the preferred source of valuation of health states, rather than the affected individuals (in contrast to #2 in Table 1), and (D) interpersonal comparisons of QALYs are permitted.

As stated in Chapter 1, health is unusual in that extra-welfarism is the dominant paradigm for economic evaluation, and CEA is preferred to CBA. Table 2 helps explain why extra-welfarism is central to QALY-based CEA. There is no reason why such principles cannot be applied to areas other than health.

Table 2: Four ways in which extra-welfarism differs from welfarism, based on Brouwer et al. (2008)

	Characteristic that differs from welfarism	**Practical consequence in QALY-based health economic evaluation**
A. Evaluative space	Permits the use of outcomes other than utility	Health *per se is* the maximand (rather than the utility derived from health). It is measured in QALYs. Other outcomes can be considered too (e.g. distribution).
B. Source of valuation	Permits valuation (weighting) based on sources other than the affected individuals	Emphasis on societal valuation based on trade-off exercises (preference elicitation) undertaken by a representative sample of citizens, but other sources permitted too (e.g. patients)
C. Basis of weighting	Permits the weighting of outcomes (whether utility or other) according to principles that need not be preference-based	More of a possibility than a central characteristic. It can be argued that, with a normative objective of fairness, more weight might be placed on health gained by the young or by people who stand to lose a large proportion of their remaining health.
D. Interpersonal comparisons	permits interpersonal comparisons of well-being* - it is contested as to whether this is permitted for utility (Arrow, 1951).	QALYs compare the health of individuals

*Brouwer et al. use “well-being” to refer to assessments of anything other than utility, and “welfare” to mean assessments of utility in the welfarist sense.

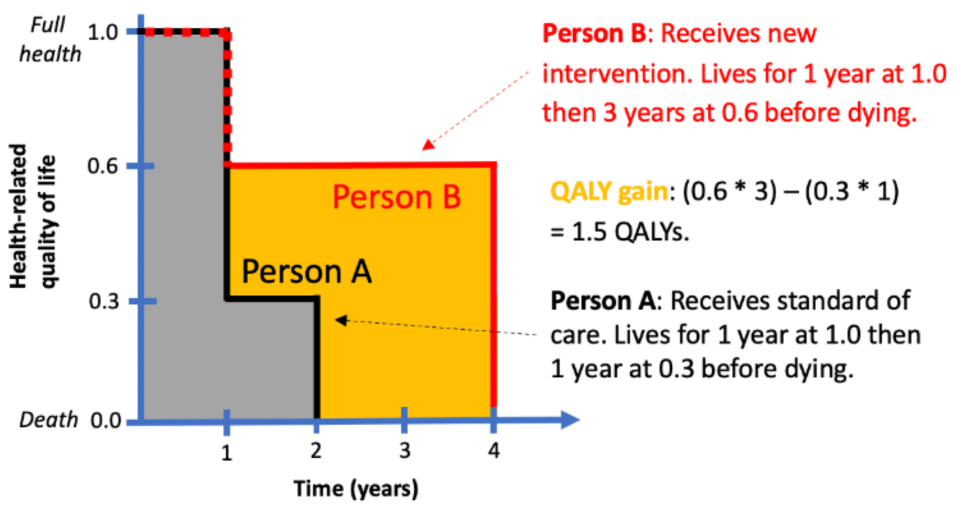
2.2.2. Measuring and valuing benefits in health

Quality-adjusted life-years (QALYs)

A quality-adjusted life year (QALY) is an extra-welfarist measure of the value of health (Drummond et al., 2015). There are extra-welfarist approaches other than QALY-based CEA, such as those using the capability approach (discussed in section 2.4). I explain QALYs and health-related quality of life (HRQoL) at length, since much of the theory will be used in later chapters.

The QALY is a type of health-adjusted life year (Gold et al., 2002), which permits the comparison of interventions for addressing diverse health issues from arthritis to Zika virus disease. QALYs combine length of life with HRQoL, which can vary during the period measured. Index values of HRQoL are measured on a 0-1 scale, anchored at 0 (dead) and 1 (full health) which is a normative decision (Brazier et al., 2016). One QALY represents one year in full health, or two years with HRQoL at 0.5, and so on (Drummond et al., 2015). Figure 1 illustrates QALYs graphically, depicting a QALY gain of 1.5 for Person B (who receives a new intervention) as compared to Person A (who receives standard of care). Note that Person B lives two years longer than Person A.

Figure 1: Visualisation of two people's QALYs over a four-year period



The source and nature of HRQoL information, and the process for its conversion into index values, are therefore important considerations when using QALYs. HRQoL is a concept which acknowledges the multi-dimensional nature of health (WHO, 1948). It is commonly measured using psychometric instruments which comprise descriptive systems of health status. These instruments usually aim to capture a person's perception of their level of health-related functioning across multiple dimensions (Fayers and Machin, 2015). Many such instruments exist. Some are generic, aiming to cover all relevant aspects of health, while others are condition-specific (Brazier et al., 2016).

The most commonly-used "preference-based" measure is the EQ-5D (Euroqol Group, 2009). The descriptive system of its five-level (5L) version of the EQ-5D is summarised in Table 3. It comprises a psychometric item for each of five dimensions of functional status (mobility, self-care, usual activities, anxiety and depression, role function). The respondent selects the level which best describes their health on that day. Based on their responses to the five questions, each of which has five levels, a person can be classified into one of 3,125 (=5^5) possible health "states". The best possible state is labelled 11111, representing the "1" level on all five dimensions (Table 3). Someone with a slight headache but no other problems might score themselves in state 11121, for example. The worst possible state is 55555.

Table 3: descriptive system of EQ-5D-5L (Euroqol Group, 2009)

<table>
<tr><th>Dimension</th><th>Framing of highest level</th><th>Five-point scale</th></tr>
<tr><td>Mobility</td><td>I have no problems in walking about</td><td rowspan="3">1 - no problems...
2 - slight problems...
3 - moderate problems...
4 - severe problems...
5 - unable to…</td></tr>
<tr><td>Self-care</td><td>I have no problems washing or dressing myself</td></tr>
<tr><td>Usual activities*</td><td>I have no problems doing my usual activities</td></tr>
<tr><td>Pain / discomfort</td><td>I have no pain or discomfort</td><td rowspan="2">1 - no/not…
2 - slight(ly)…
3 - moderate(ly)…
4 - severe(ly)…
5 - extreme(ly)…</td></tr>
<tr><td>Anxiety / depression</td><td>I am not anxious or depressed</td></tr>
</table>

*e.g. work, study, housework, family or leisure activities

Valuing QALYs

A health state can be converted into a preference-based HRQoL index value using a ‘value set’, which is an index estimated using a stated preference survey methodology (Torrance, 1986). Such stated preference methods, such as time trade-off or discrete choice experiment, ask people to trade off different states against one another by making hypothetical choices (Brazier et al., 2016). These studies are often, but not always, undertaken in a random sample of the general population such that the value set represents societal preferences over health. The use of QALYs is based on expected utility theory, involving assumptions which some have identified as questionable (Bleichrodt and Pinto, 2005; Mooney, 1989). For example, valuation studies sometimes have results which are inconsistent with expected utility theory, leading some to propose the application of other behavioural theories such as regret theory (Loomes and Sugden, 1982; Smith, 1996).

Valuation studies are also often undertaken per jurisdiction – EQ-5D value sets used in the UK are different from those used in the Netherlands. Preference-based HRQoL measures necessarily contain a small number of items (typically five or six), in order to allow stated preference methods to work properly. People are unable to trade off more than about seven attribute levels simultaneously (Hensher et al., 2015).

Regression analysis of choices made by a sample of people can be used to derive HRQoL value sets. For example, in an EQ-5D-5L value set for England, the index value of state 11112 is 0.922 while for 11121 it is 0.937 (Devlin et al., 2018), recalling that indices are anchored at 0 (death) and 1 (full health). If a person experiences a slight headache 10 days a year, and is in state 11121 during those days (slight pain or discomfort but otherwise full health), they experience 0.998 QALYs that year.[1] Recall that state 11112 has a lower index value than 11121. This shows that, on average, survey respondents’ stated preferences delivered the result that, alongside otherwise full health, “slight” anxiety or depression was worse than having “slight” pain or discomfort (see Table 3). Note that valuation done in this way is different from anxiety/depression having a higher arithmetic weight than pain/discomfort

[1] $0.998 = 1 - \left(\frac{10}{365} * (1 - 0.937)\right)$

for all possible combinations of levels. It allows for unequal intervals between levels, e.g. the interval between “slight” and “moderate” (Table 3) can be smaller in index value terms than that between “severe” and “extreme”. It also allows some health states to be valued at worse than death (Brazier et al., 2016). For example, state 55555 (see Table 3) is valued at −0.285 in the England value set (Devlin et al., 2018).

Health status is also commonly assessed using a visual analogue scale (VAS), rather than through the use of psychometric items (Drummond et al., 2015; Torrance, 1986). Use of the EQ-5D, strictly speaking, requires using a VAS after the psychometric questions (Euroqol Group, 2009). The VAS asks respondents to “Mark an X on the scale to indicate how your health is TODAY”, after explaining that the scale is numbered from 0 to 100, where 100 means the “best health you can imagine” and 0 means the “worst health you can imagine”. Such VAS scores can be used to weight QALYs as an alternative to, for example, EQ-5D index values. The use of VAS in this way is often criticised for lacking theoretical foundation and not being choice-based, though these points can be countered (Parkin and Devlin, 2006; Torrance et al., 2001). Furthermore, valuation of different dimensions of health is implicit and individual, rather than being derived from societal preferences over health states.

QALYs are routinely used in national health decision-making in the UK, the Netherlands, Australia and elsewhere (O’Donnell et al., 2009). In the UK national health service (NHS), in principle, only interventions which cost less than £20,000-£30,000 to gain one QALY should be approved, though there are exceptions (NICE, 2018). If a study finds a new drug costs £100,000 to gain a QALY then, in theory, it should not be approved. This is because approval would displace more health than it creates, by necessitating the reduction of other services elsewhere (since the budget is fixed in the medium-term). Through the threshold is arbitrary, it is often taken as the provider’s willingness to pay for a QALY. Empirical studies based on supply-side estimates of marginal productivity of the NHS suggests lower thresholds should be applied, but there remains much debate on this (Claxton et al., 2015). Similar supply-side estimates have been calculated for DALYs in LMICs (Ochalek et al., 2018).

QALYs have proven useful for priority-setting in countries which use them, but they do have several limitations (Drummond et al., 2015). Of particular importance for this research is that when interventions affect outcomes beyond health, as measured in EQ-5D and similar instruments, QALYs cannot capture these additional consequences. QALYs would not capture many of the benefits of sanitation (discussed in section 2.3 below), since they are unlikely to be reflected in HRQoL measures, as is the case for many public health interventions (Weatherly et al., 2009). This has led some to look for alternatives to HRQoL, such as more broad-based capability measures, discussed in section 2.4. There are also initiatives to build on the QALY and retain a utility framework rather than switching to the capability approach. For example, an initiative to "extend the QALY" is developing a measure of quality of life broader than the EQ-5D for use in economic evaluations across health and social care (Connell et al., 2018). Now called the "the EuroQol Health and Wellbeing instrument", it is valued on a QALY-like scale where zero is "dead" and one is "full quality of life" (rather than full health). There is also interest in using welfarist CBA for health which allows monetary valuation of health and other outcomes, especially in LMICs (Robinson et al., 2019a). Monetary valuation of QALYs is also relatively common (Ryen and Svensson, 2015).

Anchoring QALYs

QALYs are measured on an interval scale. This ensures that the difference between 1 and 2 QALYs is the same as that between 4 and 5 QALYs. However, it does not permit the conclusion that 4 QALYs has twice as much value as 2 QALYs, in the same way that 20 degrees Celsius is not "twice as hot" as 10C. However, *changes* in QALYs are on a ratio scale with a true zero, like height (Drummond et al., 2015). To value QALYs, index values for preference-based HRQoL measures such as EQ-5D and SF-6D are "anchored" at 1 = full health and 0 = dead. Practically, the anchoring process usually means rescaling data collected in preference elicitation studies, with the exact approach depending on the elicitation method such as time trade-off or DCE (Brazier et al., 2016). For cost-effectiveness modelling purposes, anchoring the "dead" state at zero is convenient because it means that people generate zero QALYs once they're dead.

Turning to the other end of the scale, the QALY approach does not assume that everyone experiences or enjoys "full health" in the same way. Rather, it takes the normative position that all individuals should have equal weight in the value set (Drummond et al., 2015). A welfarist approach based on willingness to pay, for example, would allow individual ability to pay to affect valuation at the population level.

An interval scale only requires that two arbitrary points are chosen, such as 0 and 100 on the Celsius scale. Given the methodological challenges associated with valuing the "dead" state in elicitation exercises, some have proposed means, such as the worst health state defined by the descriptive system, or as the worst health state imaginable (Sampson et al., 2020).

In the ICECAP family of capability measures (discussed in section 2.4.2 below), index values are anchored at 1 = full capability and 0 = no capability (Coast, 2019). The decision was taken, on a philosophical basis, to attach the best and worst states respectively to these anchor points (Coast et al., 2008a). This means there are no index values greater than 1 or less than 0. Death is simply a state where there is no capability, and death plays no role in anchoring. Someone can have "no capability" whether they are in a coma, dead, or simply have very low quality of life. The ICECAP measures have been valued as indices using best-worst scaling, grounded in a reading of Sen's writings that rejects the use of preferences/choices in valuation (Cookson, 2005). Rescaling raw data to anchor full capability at 1 and no capability at 0 is practically achieved by subtracting 20% of the value of the worst state from all attributes and then dividing by the index value for the best state (Coast et al., 2008a).

Disability adjusted life years (DALYs)

Disability adjusted life years (DALYs) are another extra-welfarist measure of the value of health (Gold et al., 2002; Murray and Lopez, 1996). One DALY is a one year of healthy life lost, calculated by summing years of life lost (YLL) and years lived with disability (YLD), that is: $DALY = YLL + YLD$. YLLs are estimated based on life expectancy at age of death, i.e.

years of life remaining for someone who has reached that age (Devleesschauwer et al., 2014). For cross-country comparisons, a single life table is often used so all mortality is valued equally, but for within-country comparisons life tables for that country may be used.

YLDs are measured on a scale from 0 to 1, where 0 equals a state of full health and 1 equals death, i.e. the inverse of the HRQoL index value scale. Interventions aim to gain QALYs but avert DALYs. The worse the non-fatal consequence, the higher its disability weight. For example, internationally standardised weights for mild, moderate and severe diarrhoea are 0.074, 0.188 and 0.247 respectively (GBD Network, 2018). If a 20-year old catches cholera and suffers severe diarrhoea for 5 days before recovering, the burden of disease is 0.003 DALYs.[2] However, if they die on the fifth day in a country where life expectancy at 20 is 70 years, then the burden is 50.003 DALYs.[3]

DALYs are best-known for their use in global burden of disease estimation at the population level, but they are also the most common outcome used in CEA in LMICs (Neumann et al., 2016; Pitt et al., 2016). One reason for this is the paucity of locally-collected utility data with which to estimate QALYs (Airoldi and Morton, 2009). In the assessment of interventions such as sanitation which affect outcomes not captured in disability weights, DALYs have many of the same limitations as QALYs. Disability weights are mostly focused on clinical conditions, and QoL aspects such as privacy, safety and dignity cannot be captured (GBD Network, 2018). Even sanitation-related health outcomes such as undernutrition and helminth infections are difficult to capture in DALYs (King, 2014).

2.2.3. Types of economic evaluation as applied to health

"Full" health economic evaluations include both a comparison of (i) two or more alternative interventions, and (ii) both costs and consequences, in a combined metric (Drummond et al., 2015). Other types of economic evaluation are considered "partial", such as cost analysis

[2] $0.003 = 0.247 * \frac{5}{365}$

[3] $50.003 = 50 + \left(0.247 * \frac{5}{365}\right)$

using measures such as "cost per person served". There are two main types of full economic evaluation: CEA and CBA (Table 4).

Table 4: Two main types of full economic evaluation, based on Drummond et al. (2015)

Type of study	Consequences	Measurement / valuation of consequences	Welfare economic perspective
Cost-effectiveness analysis (CEA)*	Single effect common to both alternatives (e.g. diarrhoeal disease)	Natural units (e.g. case averted)	Extra-welfarist
	Single or multiple effects, not necessarily common to both alternatives (e.g. diarrhoeal disease and acute respiratory infection)	Health-adjusted life years (e.g. QALYs, DALYs)	Extra-welfarist
Cost-benefit analysis (CBA)	Single or multiple effects, not necessarily common to both alternatives (e.g. diarrhoeal disease and travel time savings)	Monetary units (allowing calculation of a benefit-cost ratio)	Welfarist

* Some authors, especially in the UK, refer to CEA using QALYs as cost-utility analysis. I follow the international convention of referring to this as one type of CEA (Drummond et al., 2015), since DALYs are not strictly speaking a utility measure and this limits confusion.

Cost-benefit analysis

Cost-benefit analysis (CBA) is welfarist full economic evaluation (Table 4), in which costs and outcomes are valued in monetary terms (Boardman et al., 2018). CBA subtracts the discounted stream of costs from the discounted stream of benefits, to assess net present value. Other output metrics include the benefit-cost ratio (BCR) or internal rate of return. The two most important normative principles of CBA are individual sovereignty and utility maximisation (section 2.1). In theory, the monetary values placed on outcomes reflect the extent to which individuals would be willing to pay for those outcomes. Measures such as BCRs can be compared across interventions and sectors.

In general, CBA is less used in health than in other sectors (Drummond et al., 2015). This is mainly because of perceptions that such methods directly place a monetary value on human life, using methods such as value of a statistical life (VSL) (Robinson et al., 2019b). Strictly speaking, however, such methods involve valuing mortality risk reduction, rather than valuing life *per se*. Interest in CBA for global health is increasing. Initiatives such as the Copenhagen Consensus (2020) have undertaken cross-sectoral prioritisation studies, comparing BCRs for a wide range of interventions in countries including Haiti, Bangladesh, India and Ghana. A recent project established reference case guidelines for CBA in global health and development, with a view to increasing the use of the method and improving comparability of results (Robinson et al., 2019a).

Cost-effectiveness analysis (CEA)

Cost-effectiveness analysis (CEA) is a type of extra-welfarist full economic evaluation. In the health sector, outcomes are measured either in natural units (e.g. cases averted) or in health-adjusted life years such as QALYs and DALYs (Table 4). CEA can be an incremental analysis, that is, comparing costs and effects of an intervention over those of another feasible intervention such as standard of care (or added to it). The measure of comparison is an incremental cost-effectiveness ratio (ICER), namely incremental costs divided by incremental effects. However, there is increasing interest in focusing instead on net benefit, e.g. net health benefit, or net monetary benefit if the decision-maker's willingness to pay threshold is known (Drummond et al., 2015). This is because of some unhelpful statistical properties of ICERs as ratios, which precludes certain econometric analysis.

CEA can also be used as an average analysis, that is, comparing to 'no intervention', generating average cost-effectiveness ratios (ACERs). Average analyses can be useful for providing an overall picture of allocative efficiency within a sector (Hutubessy et al., 2003; Murray et al., 2000). However, for supporting decisions between options competing for the same resources, incremental analyses provide the only realistic understanding of opportunity costs (Hoch and Dewa, 2008). The use of ACERs in is therefore generally strongly discouraged by health economists.

When natural units are used, CEA can only assess a single effect common to both alternatives, and compare metrics such as "cost per diarrhoea case averted". QALYs and DALYs, however, can incorporate multiple effects of an intervention, e.g. on both diarrhoeal disease and acute respiratory infections. Furthermore, these do not have to be common to both alternatives. This facilitates comparison of multiple interventions within the health sector. The “cost per DALY averted” can be compared between interventions against malaria and against cataracts.

Cost-consequences analysis and multi-criteria decision analysis

Two more types of economic evaluation are important to mention. Cost-consequences analysis (CCA) and multi-criteria decision analysis (MCDA). Cost-consequences analysis is a type of partial economic evaluation in which costs and various outcomes are tabulated in a disaggregated balance sheet (Mauskopf et al., 1998). In addition to health outcomes, outcomes can be within health (e.g. patient satisfaction) or beyond health (e.g. well-being measures, criminal justice outcomes). The decision-maker must then form their own weighting of those bits of information according to their perceived relevance and importance (Drummond et al., 2015).

This has the downside of there being no transparent decision rule, like thresholds for QALYs and DALYs. This leaves too much to the discretion of the decision-maker, which may lead to cherry-picking (Hunter and Shearer, 2014). Nonetheless, it may allow equity to come into decision-making more explicitly, and allows transparent presentation to decision-makers of disparate pieces of qualitative and quantitative information (Coast, 2004). CCA has been recommended for assessing public health interventions as an initial analytic step before selecting a full economic evaluation method (Weatherly et al., 2009). It is also included in the reference case of the National Institute of Health and Social Care Excellence (NICE) as permitted evidence for decision-making regarding interventions funded by the public sector with health and non-health outcomes (NICE, 2020).

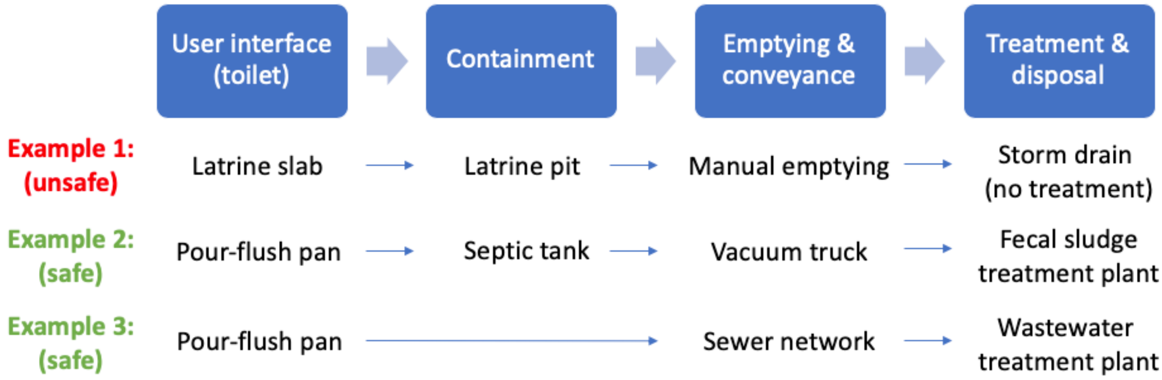
Multi-criteria decision analysis (MCDA) assesses options against a number of criteria, making the impact of the decision of all the criteria explicit, and making the relative importance of the criteria explicit (Devlin and Sussex, 2011). In other words, it is a CCA with numerical scores for all outcomes, and explicit and transparent weights. Weights are subjective, though the analysis can be transparent about the process used to identify them, and data for some scores can be subjective too. Some have advocated for more use of MCDA in health, arguing that in real life many criteria inform a decision, not only an incremental cost-effectiveness ratio or benefit-cost ratio, and it would be better to be transparent about the relative importance given to these factors (Baltussen and Niessen, 2006; Devlin and Sussex, 2011). MCDA was identified as one possible solution to the challenge of economic evaluation of public health interventions (Weatherly et al., 2009). It is not explicitly part of the NICE (2020) reference case but the document does note that MCDA could be a tool in a CCA process. Furthermore, it can be argued that QALYs are in fact a form of MCDA in which different dimensions of health are transparently weighted and traded-off with length of life (Drummond et al., 2015).

2.3. Economic evaluation as applied to sanitation

2.3.1. What are sanitation services?

Sanitation is the access to and use of facilities and services for the safe disposal of human urine and faeces (WHO, 2018). Implicit in the “use” element to the definition is that sanitation is about behaviour, i.e. how people manage their practices of defecation, urination, practice menstrual hygiene, and so on. Safe sanitation systems ensure the separation of human excreta from human contact through the “sanitation service chain” (Figure 2). A toilet is the user interface with a sanitation system, e.g. a latrine slab, pour-flush pan, or pedestal seat where excreta are captured. A safe system might also employ sewers, vacuum trucks, and treatment plants, or in a rural area might simply mean safely sealing the latrine pit when full.

Figure 2: Sanitation service chain, based on WHO (2018)



The sustainable development goal (SDG) indicator for sanitation (6.2.1) measures "safely managed" sanitation services (UNICEF & WHO, 2019). Its definition incorporates minimum quality standards for both the toilet and safe management of waste along the service chain. There are a large number of safe and unsafe routes excreta can take down the service chain - Figure 2 provides some illustrative examples. The step below safely-managed in the SDG level of service framework is "basic" sanitation, which considers only whether toilets are "improved" and the extent to which they are shared. Approximately 2 billion people globally lack access to a "basic" level of sanitation service (UNICEF & WHO, 2019). In other words, a quarter of humanity either practices open defecation (9%), uses a toilet which is unimproved (9%), or one which is improved but shared with other households (8%).

In what follows, I focus predominantly on urban sanitation (rather than rural) since urban areas are the focus of this thesis. Of the people without basic sanitation globally, 30% are in urban areas, but the urban challenge is growing rapidly (UNICEF & WHO, 2019). Around a billion urban dwellers live in overcrowded tenements, informal settlements or temporary camps characterised by a lack of basic services (Mitlin and Satterthwaite, 2013). By 2030, one in three people globally will live in a city of at least 500,000 inhabitants (United Nations, 2016a). Sub-Saharan Africa's urban population is predicted to double between 2015 and 2035 (United Nations, 2017). As LMIC populations increasingly urbanise, the sanitation challenge will too.

2.3.2. Sanitation as a state of being

The above section focused on sanitation as a service and technology. It is useful to classify the scale of a problem in terms of technology or service standards (such as those in SDG 6), in order to be able to compare across settings. For example, to discuss interventions and their effectiveness, it is important to be able to characterise the 'before' and 'after' situations of people benefitting from interventions. However, the evaluation of specific interventions would ideally include a more nuanced understanding of changes brought about in people's lives. Just because two people have achieved the same service level improvement, it does not follow that their perceptions of sanitation-related privacy have been improved identically, or even at all.

Sanitation, then, can be considered a state of being. Having "access to and use of" health services, to paraphrase the WHO (2018) definition of sanitation, is generally understood as being different from having health itself. I would argue that, in the same way that someone has a self-perceived level of health, they also have a self-perceived level of sanitation. It is a reasonable assumption that using a better toilet technology might be associated with a higher self-perceived level of sanitation. However, this assumption has not been tested in formal studies. Very few psychometric measures related to sanitation exist, as will be discussed below, and none are regularly used. Health is defined by WHO (1948) as "a state of complete physical, mental, and social well-being". An analogous definition for sanitation as a state of being does not exist. One study has defined "sanitation insecurity" (Caruso et al., 2017b), by undertaking qualitative work and building on an earlier conceptualisation of "toilet insecurity" (O'Reilly, 2016). Based on qualitative research in rural India, Caruso et al. defined sanitation insecurity as:

> *"Insufficient and uncertain access to a socio-cultural and social environments that respect and respond to the sanitation needs of individuals, and to adequate physical spaces and resources for independently, comfortably, safely, hygienically, and privately urinating, defecating, and managing menses with dignity at any time of day or year as needs arise in a manner that prevents fecal contamination of the environment and promotes health."* (Caruso et al., 2017b)

This definition centres on “needs”, but the study does not refer to a normative framework for needs. Options might include the concepts of the hierarchy of needs (Maslow, 1943), basic needs (Streeten, 1982) or human needs (Doyal and Gough, 1984). One option for conceptualising sanitation as a state of being is to adapt, and shorten, the sanitation insecurity definition. An individual’s “sanitation state” might be defined as “the extent to which an individual’s needs are met with respect to urination, defecation, and menstrual hygiene through access to social environments and physical resources.” Below I propose an alternative conceptualisation of sanitation states, based on the capability approach (Sen, 1980), which addresses this.

2.3.3. Characteristics of an individual’s sanitation state

In health, the characteristics of health states are implicitly defined by dimensions in the psychometric measures used to weight QALYs. In the EQ-5D (section 2.2), these characteristics are the extent of problems related to mobility, the extent of pain or discomfort, and so on (Table 3). What characteristics, then, might define an individual’s sanitation state? A systematic review of the relationship between sanitation and mental well-being identified privacy and safety as root dimensions, alongside the related concepts of shame, anxiety, fear, assault, dignity and embarrassment (Sclar et al., 2018). Beyond mental well-being, cleanliness and convenience are also commonly reported as important by users (Novotný et al., 2018). Studies have explored the value of sanitation to users from different disciplinary perspectives. Several studies focus on motives, defined as “mechanisms designed by evolution to cause animals to seek to meet a need through behavior” (Aunger and Curtis, 2013). Motive-oriented studies aim to identify the drivers of a behaviour, such as practicing open defecation (OD), often with disciplinary roots in psychology (Jenkins and Curtis, 2005; Jenkins and Scott, 2007; Mukherjee, 2001).

Another group of studies aims to identify sources of insecurity and stress as factors affecting mental wellbeing, with a public health orientation (Caruso et al., 2017b; Kwiringira et al., 2014; Sahoo et al., 2015; Shiras et al., 2018). Stress is defined as the result of real or perceived threats that exceed a person's ability to manage them (Shiras et al., 2018). Sanitation insecurity was defined above, and a psychometric measure of sanitation insecurity amongst women comprising seven factor scales has been developed (Caruso et al., 2017a). It has been used in evaluative studies as both an exposure (Caruso et al., 2018) and as an effect moderator (Delea et al., 2019) on the causal pathway to outcomes such as mental well-being. The authors therefore see women's sanitation insecurity as an exposure not an outcome. To my knowledge, this is the only measure of some aspect of sanitation states. However, other scales exist for different aspects of sanitation, such as one assessing attitudinal determinants of sanitation uptake and use (Dreibelbis et al., 2015). Visual analogue scales have, to my knowledge, not been used in the valuation of sanitation states.

It is important to note emerging efforts to measure empowerment in WASH (Caruso et al., 2020; Dickin et al., 2021), which are predominantly focused on processes rather than QoL outcomes. For example, the Empowerment in WASH Index focuses on aspects like roles and responsibilities, decision-making processes, and leadership (Dickin et al., 2021). The urban sanitation empowerment scale aims to focus on similar aspects, but also on knowledge, skills and norms, as well as more outcome-oriented dimensions of bodily integrity and security (Caruso et al., 2020). Neither uses the capability approach.

2.3.4. Sanitation interventions

There is no standard typology of sanitation interventions. The SDG 6 framework specifies technologies (e.g. piped water) at levels of service (e.g. on-plot, continuous). However, there are many types of interventions which could result in those technologies or levels of service, either in new access or in improving the quality of existing access. A simplification for interventions targeting households is to say that most aim to change individuals' behaviours, either by providing assets/subsidies, by promotion alone, or a combination of both (WHO, 2018). A recent systematic review investigated the impact of sanitation interventions on latrine coverage and use, focusing only on the toilet stage of the chain (Garn et al., 2017). The review categorised interventions as follows: (i) latrine subsidy/provision, (ii) sewerage, (iii) sanitation education, (iv) community-led total sanitation (CLTS). In practice, a project or programme often include several of these, e.g.

CLTS with subsidy for the poorest, or latrine provision with education. Beyond toilets, investments aiming to improve pit emptying services or constructing a treatment plant are also sanitation interventions.

All toilet construction involves engineering, in the sense that the design, soil structure, water table, and materials used will affect the structure's quality and durability. Building anything more advanced than a basic pit latrine requires some level of construction expertise. Public health practitioners are required as well (Luby, 2017). The sanitation sector therefore straddles the realms of public health and public works. In most LMICs, sanitation investments are rarely funded and/or directed by the Ministry of Health. Rather, this is the role of the Ministry of Public Works, the Ministry of Water, or a shared responsibility between several ministries (WHO, 2019). In urban settings utilities also often play an important role, not only in managing sewer networks but in managing and regulating on-site sanitation and faecal sludge management services as well (Scott et al., 2019).

Interventions based on promotion or campaigns do receive public funds (De Buck et al., 2017; Dreibelbis et al., 2013). However, the vast majority of public finance for sanitation in LMICs is spent on delivering public health outcomes through engineering means (WHO, 2017), particularly in urban areas. Often this involves constructing public infrastructure, but sometimes private assets too. The most successful sanitation campaign in recent history is the Swachh Bharat Mission - Gramin in India (Curtis, 2019). Of the US$ 5.2 billion of public expenditure on this programme between 2014-17, 97% was on ex post subsidy incentives for households to construct private toilets, and 1% on the information, education and communication component (Mehta, 2018). However, the focus of Swachh Bharat was overwhelmingly (though not exclusively) rural. Urban settings present different challenges, and municipal authorities are the level of institution most likely to be making practical investment decisions. Urban sanitation interventions are rarely likely to involve funding private household toilets. The kinds of interventions municipalities might fund, at the level of containment, are public toilet facilities, incentives for improving shared household sanitation, or promotional campaigns encouraging upgrading of private household latrines. The lion's share of funds for urban sanitation are likely to be spent on public infrastructure for the rest of the service chain (Figure 2), for example sewer networks or faecal sludge management (FSM) services. Municipalities might also undertake interventions to promote household uptake of such services. A further body of possible interventions relates to coordination and regulation of private providers (such as vacuum tanker operators).

2.3.5. Measuring and valuing sanitation-related benefits in economic evaluation

A recent study tabulates many economic evaluations of WASH interventions (Hutton and Chase, 2016). However, it was a broad evidence review which was not systematic, and study quality was not assessed. Moreover, economic evaluation was only briefly incorporated into a wide-ranging narrative, discussing WASH services as a whole rather than the relative merits of specific intervention types. There has been no systematic review of economic evaluations to date for sanitation interventions, which was the primary rationale for undertaking the study reported in Chapter 4. Below, I focus on the methods for benefit valuation and the extent to which quality of life benefits have been measured and valued, saving a detailed discussion of the empirical literature for Chapter 4.

Types of benefits of sanitation

Table 5 presents the types of outcome which have been measured and valued in CBA studies of sanitation interventions. It is separated into those which have frequently been included, versus those which have not typically been quantified, or have been measured but not in such a way as can be included in headline results. A description is provided for each, as well as valuation methods which have been used. An annex to the systematic review (Chapter 4) sets out which benefits were measured and valued in each of the studies included in the review and their relative contribution to total benefits. Notably, it was very rare for any benefits to be valued via stated preference willingness to pay studies, though some estimates ultimately derived from revealed preference studies (e.g. value of statistical life, value of time).

Table 5: Types of outcome valued in economic evaluation of sanitation interventions (based on Hutton (2020), Whittington et al. (2009), Hutton (2014))

Impact area	Type of outcome	Description	Common valuation methods
Outcomes frequently valued and included in headline results			
Averted direct health	Mortality	The value of mortality risk reductions, predominantly from diarrhoea	Value of a statistical life (VSL), human capital approach (HCA), or 'years of life lost' (YLL) component of DALY (Robinson et al. 2019b)
	Morbidity - cost of illness (COI) to the patient	Direct costs borne by patients/families (e.g. transport, fees, medicines), from diarrhoea but also helminth infections and other diseases	Willingness to pay to avoid disease, financial expenditure or 'years lost due to disability' (YLD) component of DALY (Robinson & Hammitt, 2018)
	Morbidity - COI to the health system	Direct costs of treating patients (e.g. clinicians' time)	Opportunity cost of clinicians' time (Robinson & Hammitt 2018)
Averted indirect health	Lost time of patients	Lost wages, schooling or opportunity cost of time	Value of time, usually taken as a proportion of an appropriate wage (Whittington & Cook, 2018)
	Lost time of caregivers	Caring for the sick, e.g. lost wages or opportunity cost of time	
	Future productivity of children	Impact of future earnings caused by stunting, or missing school days	Not commonly included. This typically involves an assumption of impact on future wages, e.g. that a z-score increase in height for age increases future wage by 8% (Dickinson et al., 2015), or that missing a small number of school days has a linear effect on future wages (Sklar, 2017). Neither of these methods is particularly robust.
	Health externality	health gains to individuals due to the sanitation choices of others in the community.	Included as an additional reduction in diarrhoeal disease by Radin et al. (2020) and Hutton et al. (2020), based on the Andres et al. (2017) finding in rural India that community coverage above 75% is associated with greater reductions in diarrhoea. Such an externality has not yet been quantified for urban areas, where the safe containment, conveyance and treatment of faecal waste is likely to be more important than reducing open defecation.
Travel time savings	Time savings or convenience	Time saved when people start using household toilets instead of OD sites or public toilets.	Value of time, usually taken as a proportion of a wage (Whittington & Cook, 2018)
Other	Other coping costs	Financial costs averted as a result of the intervention, such as avoided fees for public toilet use.	Financial cost of the coping mechanism. Any time savings would be captured above.
	Reuse	Fertiliser replacement through use of urine and compost from faeces.	The opportunity cost of purchasing fertiliser to achieve the equivalent nutrient value (Dasgupta et al., 2019)
Outcomes typically not quantified, or not included in headline results			

Impact area	Type of outcome	Description	Common valuation methods
Quality of life	Users' quality of life directly related to service use	Including safety, dignity, convenience, social status, privacy and avoided disgust	Never yet directly valued , though sometimes asked about in surveys. For example, survey respondents are sometimes asked about their levels of satisfaction with different aspects of sanitation (Hutton et al., 2014) or their extent of agreement with statements about benefits of owning a toilet (Hutton et al., 2020).
	Clean neighbourhood externality	The welfare gain from living in a cleaner neighbourhood being more pleasant	Never valued, though sometimes asked about in surveys (Hutton et al., 2014)
Property value	Property value	The value of a property being increased by the presence or improvement of a toilet	Rarely measured. Hutton (2020) includes it as a one-off benefit at the end of the toilet's useful life, by asking households what they think the increase in property value would be. This is likely to be quite prone to bias.
Tourism	Economic benefits of increased tourism	Increased tourism revenues when destinations (e.g. beaches, parks) are more attractive due to cleaner water or environs	Never valued, though sometimes asked about in surveys (Hutton et al., 2014)

Health benefits have long been a key part of the rationale for public investment in sanitation, as mentioned in Chapter 1 (Bartram and Cairncross, 2010; Briscoe, 1987; Churchill et al., 1987). A recent meta-analysis estimated that improved sanitation reduced childhood diarrhoeal disease by 25% (95% CI: 12-37%) as compared to unimproved sanitation (Wolf et al., 2018). It also suggested that larger effects were achieved with higher levels of service (i.e. sewerage versus pit latrines). However, several recent randomised trials found little or no effect of sanitation interventions on disease (Humphrey et al., 2019; Knee et al., 2020; Luby et al., 2018; Null et al., 2018). Those trials looked not only at diarrhoea but also stunting and environmental enteropathy. The consensus amongst epidemiologists active in this field is that the case for sanitation remains strong, but that basic services alone are unlikely to have a large short-term impact (Cumming et al., 2019). When comparing this evidence base to that of health interventions, it is important to note that the systematic review evidence for the impact of sanitation on diarrhoea only really draws conclusions for broad levels of service rather than specific interventions. There are many ways to encourage uptake of improved sanitation (Garn et al., 2017). For economic evaluation purposes this means that, except for the very few trial-based studies (discussed in Chapter 3), almost all studies apply the same risk reduction assumption from the systematic review literature.

As noted in Table 5, some of these benefits have rarely or never been measured or valued in sanitation CBA studies. These exclusions may result in misallocation of resources if interventions differ in the extent to which they produce these benefits, as appears likely. The focus of this thesis is on making headway on measuring and valuing the quality of life benefits, particularly those experienced in relation to using the toilet (rather than the gain from living in a cleaner neighbourhood). Including these QoL outcomes in economic evaluations would support allocative efficiency. The exclusion of QoL outcomes is particularly problematic as these are often cited by toilet users as the most important perceived benefits or motives for investment (Jenkins and Curtis, 2005; Jenkins and Scott, 2007; Mukherjee, 2001). As such, they are likely to underpin household willingness to pay for sanitation improvements, and thus be important in welfarist CBA.

Focus on quality of life benefits of sanitation

A systematic review of motivations and contextual factors affecting rural sanitation in LMICs reviewed 40 studies from 16 countries (Novotný et al., 2018). While the review was focused on rural areas, one of its results tables is reproduced below (Table 6) to illustrate the wide range of topics which might arise in discussions with people about their sanitation behaviours and practices. Concepts related to QoL which were identified in the review as arising with more than 10 observations are: (i) convenience / time savings, (ii) privacy, (iii) safety, (iv) smells and insects, (v) prestige and status, (vi) health, and (vii) cleanliness.

Table 6: "Contextual factors and motivations" related to sanitation, reproduced from Novotny et al. (2018)

Broader types of motivations and factors (N = number of observations; S = number of studies; C = number of countries)	**More specific types of motivations and factors** (number of observations)
Privacy, safety, convenience (N = 107; S = 25; C = 14)	Convenience of latrine (30); Privacy (22); Safety (16); Time savings (10); Sharing latrine (8); Smell from latrine (8); Convenience of open defecation (4); Latrine attracts insects (3); Unsafe latrine (3); Smell from OD (3)
Socioeconomic factors (N = 99; S = 34; C = 16)	Wealth (25); Housing quality (10); Income (8); Ownership of livestock (1) Lack of money (14); Cost of latrine (or its perception) (13); Access to loan (individual) (2) Education (general) (13) Non-agricultural occupation (9); Nomadic lifestyle (2); Agricultural occupation (1); Low status occupation (1)
Sanitation infrastructure, maintenance, supply, access to materials or manpower (N = 79; S = 27; C = 15)	Acceptable quality of latrine (16); Low quality, incomplete, or broken latrine (10) Availability of material (11); Lack of manpower (9) Available sanitation supply services (13); Lack of skills or technology support (2); Information on supply-side options (5) Need for latrine maintenance (13)
Spatial and environmental factors (N = 79; S = 29; C = 15)	Access to water (for use) (18); Soil, bedrock, suitable terrain for latrine (13); Sufficient space for OD (13); Location (centrality, accessibility, etc.) (10); Climate constraints (flooding, rain, etc.) (9); Lack of space for latrine construction (9); Nearby river, lake, swamps (7)
Social pressure, networks, and learning (N = 77; S = 28; C = 14)	Prestige, status (21); Social pressure (12); Effort to be modern (5); Shame prevention (2) Social networks (20); Social capital (11); Social learning (6)
Health and/or cleanliness (N = 51; S = 23; C = 14)	Health-related expectations (28); Cleanliness (18); Health problems (experienced) (5)
Demographic characteristics (N = 34; S = 19; C = 10)	Age (9); Female head of household (7); Village size (4); Household size (4); Presence of women (3); Ethnicity (2); In-migration (2); Presence of children; (2) Presence of men (1);
Institutional support and availability of general infrastructure (N = 27; S = 15; C = 10)	Institutional support (7); Infrastructure and social services (general) (5); Subsidies (5); Involvement of local leaders or officials (5); Institutional pressure, command, sanctions (3); Access to loan (2)
Hygiene and sanitation knowledge, experience, habits (N = 21; S = 12; C = 11)	Knowledge of hygiene and sanitation advantages (14); Experience with latrine (3); Prevalent practice of open defecation (2); Soap usage (1); Hand washing (1)
Satisfied, other priorities (N = 17; S = 11; C = 6)	Satisfied with current practice (10); Other priorities (7)
Cultural factors (bylaws, taboos, etc.) (N = 13; S = 9; C = 10)	Cultural factors (bylaws, taboos, etc.) (12); Religion (Muslims) (1)
Other (N = 9; S = 7; C = 5)	Workplace far from latrine (4); Children's' playground far from latrine (1); Latrine used as storage (1); Rented house (1); Reuse of excreta (1); To increase house rent (1)

To inform my thinking on framing the relative importance of user-reported benefits, I collated studies which had implicitly or explicitly asked responded to rank benefits, motives or stressors related to sanitation. This also served to inform methods applied in the qualitative study in Chapter 5. This was a non-systematic search based on available systematic reviews (Novotný et al., 2018; Sclar et al., 2018), and forward-searching citations of identified the key references in Google Scholar. Table 7 presents the seven identified studies which met the ranking criterion, as well as presenting sufficient transparent methodological detail on what respondents were asked and how (Gross and Günther, 2014; Hulland et al., 2015; Jenkins, 1999; Jenkins and Curtis, 2005; Jenkins and Scott, 2007; Kiyu and Hardin, 1993; Mukherjee, 2001).

Table 7: Studies with rankings of sanitation benefits or motives

Study	Method / sample	Nature of question / data	Advantages / benefits / demand drivers (top 5, if >5)	Comments
Kiyu & Hardin (1993)	Quantitative - 835 women in rural Malaysia	Reasons for using a latrine for defecation, amongst regular users (frequency of response)	1. keep the compound clean 2. convenient to use 3. to keep the compound from being smelly 4. prevent the spread of illness	Appears response categories pre-specified
Jenkins (1999)	Quantitative - 320 heads of households from six villages in rural Benin (80% male)	Most important advantage of latrine adoption (respondents asked to pick ordered 3 from list)	1. avoid discomforts of the bush 2. gain prestige from visitors 3. for health 4. have more privacy to defecate 5. easier to defecate (if age/sickness)	Of the tables Jenkins presents, this is the best data to compare to other studies, since 'for health' was not included as a pre-specified category and only included spontaneously
Mukherjee (2001)	Qualitative - 36 focus groups with women and men who had built latrines, in rural Cambodia, Indonesia and Vietnam,	Benefits from having household latrines (unclear how administered)	1. clean home free of smell and flies 2. convenient day and night / time-saving 3. disease reduction 4. saving medical costs and lost wages 5. safety	Unclear if list pre-specified or all responses autonomous
Jenkins & Curtis (2005)	Qualitative - interviews with 40 heads of household in rural Benin (63% latrine adopters, 83% male)	Open-ended, including the topics (i) decision to install a latrine, and (ii) pros/cons of latrines	1= Express new experiences and lifestyle 1= Family health & safety 2= Affiliate with urban elite 2= Convenience & comfort 5. Protection from supernatural threats	No pre-specification. Health is in the broad sense. Ranking is by 'frequency of drives expressed'
Jenkins & Scott (2007)	Quantitative - nationally-representative sample of 399 rural and peri-urban households not yet adopting toilet (100% female)	Top three reasons for building, with attributes as prompted list	1. Convenience 2. easy to keep clean 3. good health 4. cleanliness 5. visitors/guests	Unclear if list pre-specified or all responses autonomous and interpreted by enumerators

Study	Method / sample	Nature of question / data	Advantages / benefits / demand drivers (top 5, if >5)	Comments
Gross & Günther (2014)	Quantitative - 2,000 respondents in rural Benin (70% female)	Motivational factors for latrine construction (appears to be presented as agree / disagree)	1. Danger and Security (night, animals) 2. Avoid diseases 3. Facilitate defecation / time savings 4. Maintain cleanliness around the house 5. Avoid seeing excrement of others	Pre-specified list
Hulland (2015)	Qualitative – 60 respondents in rural and urban Odisha, India (100% female)	Stressors commonly encountered when practicing sanitation	1. Rain / Night / Animals (safety) 2. Health 3. Encountering ghosts 4. Being seen 5. Reputation	Pre-specified list from earlier in-depth interviews

All seven studies were in rural areas, though a third of the Hulland et al. (2015) sample was reported as being urban. Of the tabulated studies, three were qualitative and four quantitative. Most explicitly or implicitly focused on motives for transitioning from open defecation (OD) to use of latrines. Rankings of benefits were not consistent, which is not surprising since individual and environmental factors affect these outcomes, and study methods varied. Health was usually framed in terms of infectious disease, or only as "health", rather than the broader well-being framing in the WHO (1948) definition. Health or infectious disease was predominantly in second or third position in rankings. Other benefits most often ranked in the top three were cleanliness or avoiding disgust, status or pride, and convenience. There are several possible explanations for variation in studies' results. Some interviewed people already with toilets, some without. Studies used different elicitation methods – some asked people to agree or disagree with a pre-specified list, while others used unprompted elicitation techniques. Some studies used qualitative methods and some quantitative.

The conclusion emerging from these studies is that rural toilet users see disease prevention as one of the benefits of moving from OD to using a basic pit latrine. However, disease prevention was often not seen as important as other benefits of sanitation, such as cleanliness or avoiding disgust, status or pride, and convenience. In the absence of urban studies, it is unknown whether such findings would extend to urban areas, or to service level transitions other than OD to pit latrines.

2.4. The relevance of the capability approach to sanitation

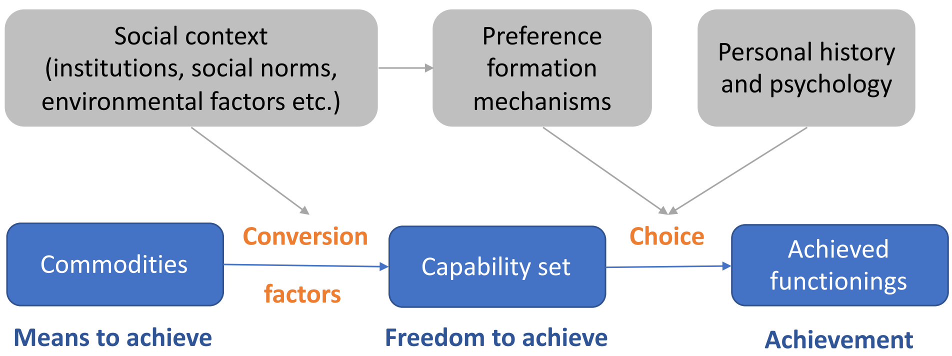
The capability approach is a way of thinking about welfare. This next section gives an overview of what it is, and then how it has been applied in health and in sanitation to date.

2.4.1. What is the capability approach?

Amartya Sen developed the capability approach over a period of years, originating in his critiques of welfarism and utilitarianism (Sen, 1980). The approach was further elaborated in subsequent works (Sen, 1993, 1985), and popularised in a more accessible book "Development as Freedom" (Sen, 1999). Others have contributed substantially to the development of the capability approach, in particular Martha Nussbaum (2011). The approach has been characterised as a conceptual framework for a range of evaluative exercises, primarily the assessment of individual wellbeing, but also the evaluation and design of policies and institutions (Robeyns, 2017).

Sen (1985) argues that commodities do not have an intrinsic benefit, but have value through allowing people to be or do things. That is, commodities can bring about capabilities. However, people have freedom to choose between capabilities in order to actually achieve functionings. Sen's classic example is the difference between someone fasting, who is choosing not to eat, and someone starving. The focus is not on material well-being but the opportunity to achieve well-being. The distinction between commodities, capabilities and functionings can be demonstrated graphically as in Figure 3 (Robeyns, 2005).

Figure 3: A person's capability set within social and personal context, adapted from Robeyns (2005)



An example to illustrate the diagram is a bicycle as a commodity providing a transportation capability, which a person can choose to take up or not. The bicycle as a commodity provides no benefit in well-being terms if it is left to rust. Furthermore, the personal characteristics of the individual affect both their capabilities and their conversion factors, as does their social and physical environment. Someone who cannot use their legs will get no benefit from a bicycle. Someone who lives in a mountainous area with poor roads may find it impossible to practically use a bicycle. Someone who lives in a dangerous city may prefer not to risk being mugged while on a bicycle. This example serves to show that the bicycle itself is not a sufficient condition for someone to have a capability or achieve a functioning.

Capabilities, then, are "beings and doings". They are the set from which someone can choose their functionings, which are what is actually achieved. Choice is important because it implies agency, another important concept in the capability approach. Agency represents the opportunity to achieve well-being, but people also have goals which may involve achievements beyond improving their own well-being (Sen, 1993). It is often said that the capability approach is underspecified, primarily because it is a normative theory (Robeyns, 2017). It is a framework for conceptualising well-being rather than explaining it. Nussbaum (2011) developed a list of ten "central human capabilities", such as life, bodily health and bodily integrity. Sen, meanwhile, has resisted ever providing such a list, arguing that public deliberation rather than expert deduction is the right way to identify what is valuable (Sen, 2005).

2.4.2. How has the capability approach been applied in health economics?

Health capability within health economics

From a philosophical perspective, a major development in applying the capability approach to health is the work of Venkatapuram (2011) on the concept of health justice. Amongst health economists, it has long been debated how the capability approach can be applied in economic evaluation (Coast et al., 2008c, 2008b). Some have explicitly used the capability approach in considering extra-welfarism in general (Culyer, 1989). Some have argued for reinterpretation of the QALY as an index of the value of a capability set (Cookson, 2005). The case is made on the basis that the descriptive systems of most preference-based measures are measuring functionings, as is the case for the EQ-5D (Table 3). The policy-maker's objective would become maximising the sum of individual capability sets. This approach has not caught on, with three main critiques. Coast (2009) takes issue with the objective of maximisation *per se*, and elsewhere has proposed focusing instead on what it would take for everyone to reach a "decent minimum" (Coast et al., 2008b). This was followed by the development of the concept of "sufficient" capability, discussed further below (Goranitis et al., 2017; Mitchell et al., 2015). Second, capability advocates within health economics tend to argue that capturing non-health sources of value is a key benefit of the capability approach in the first place (Coast et al., 2008c; Greco et al., 2016). HRQoL measures like the EQ-5D (and by extension QALYs) are too narrowly focused on health, and the same applies to DALYs. Third, there is evidence that that capability measures capture different things to HRQoL measures. Studies have explored this empirically by assessing correlations between capability-based outcome measures (see next section) and the EQ-5D, concluding that QALYs are unlikely to be a reliable proxy of capability well-being (Franklin et al., 2018; Mitchell et al., 2017).

Measuring and valuing health capability

Many health economists working on capabilities have found a middle ground between rejecting QALY-like approaches and reinterpreting the QALY. In most cases, this has involved developing capability-based outcome measures which go beyond health, but still aim for practical use in health-related economic evaluation in similar ways to HRQoL measures (Coast et al., 2015). The nine such measures developed to date are listed in Table 8. In terms of valuation, the ICECAP family has predominantly used best-worst scaling (Coast et al., 2008a; Flynn et al., 2007). Most other measures in Table 8 use equal weighting, while Greco (2016) compares different weighting options. Part of the motivation for moving towards the capability approach, particularly in the UK context, appears to also have been the need to improve economic evaluation of social care, and to undertake combined evaluation of social care outcomes alongside health (Grewal et al., 2006; Netten et al., 2012).

Table 8: Capability measures developed for the purpose of health-related economic evaluation

Family	Outcome measure	Intended purpose	Items and levels	Key reference
ICECAP* (ICEPOP Capability)	ICECAP for Adults (ICECAP-A)	Economic evaluation of health and social care for all adults	5 items, each with 4 levels	Al-Janabi et al. (2012)
	ICECAP for older people (ICECAP-O)	As ICECAP-A but for older people (developed with 65+ sample)	5 items, each with 4 levels	Grewal et al. (2006)
	ICECAP Supportive Care Measure (ICECAP-SCM)	Economic evaluation of end-of-life care	7 items, each with 4 levels	Sutton & Coast (2014)
	ICECAP Close Person Measure (ICECAP-CPM)	Economic evaluation of end-of-life care focusing on people close to the patient	6 items, each with 5 levels	Canaway et al. (2017)
Oxford capability measures (OxCAP)	OCAP-18	Economic evaluation of public health interventions	18 items	Lorgelly et al. (2015)
	OxCAP for mental health (OxCAP-MH)	Economic evaluation of mental health interventions	16 items, each with 5 levels	Simon et al. (2013)
Other	Adult Social Care Outcomes Toolkit (ASCOT)	Economic evaluation of social care-related quality of life amongst adults	8 items, each with 4 levels	Netten et al. (2012)
	Women's Capabilities Index (WCI)	Economic evaluation of women's groups in Malawi and Uganda	70 variables in 6 dimensions, varying levels.	Greco et al. (2015)
	Chronic pain measure	Economic evaluation of treatments for reducing chronic pain	8 items, each with 4 levels	Kinghorn et al. (2014)

*the Carer Experience Scale is part of the ICECAP family but not capability-based (Al-Janabi et al., 2008)

The ICECAP family of measures appear to have gained the most use so far. The ICECAP team are working on measures for under-18s to allow assessment of capabilities across the life course (Coast, 2019). The ICECAP-A (for adults) has been most widely used, and has been translated into 10 languages. Its descriptive system is summarised in Table 9 – it implies states ranging from 11111 (worst) to 44444 (best), the opposite way around to the EQ-5D and with only four levels. Note the focus on "able to X" and "can X" in the response categories. This explicitly focuses on capabilities as opposed to functionings as used in the EQ-5D (Table 3). Some have criticised these formulations for having unnatural language in some cases or bringing the illusion of semantic nuance only (Sampson, 2017). The developers of the ICECAP-A explored the impact of the language used by comparing the results of a “functionings” version to the standard version (Table 9), finding some evidence for differences between capabilities and functionings (Al-Janabi, 2018).

Table 9: descriptive system of ICECAP-A (Al-Janabi et al., 2012)

Dimension	Framing of highest level	Four-point scale
Feeling settled and secure	I am able to feel settled and secure in all areas of my life	4 - able to … in all 3 - able to … in many 2 - able to … in a few 1 - unable to … in any
Love, friendship and support	I can have a lot of love, friendship and support	4 - can have a lot 3 - can have quite a lot 2 - can have a little 1 - cannot have any
Being independent	I am able to be completely independent	4 - able to be completely... 3 - able to be … in many things 2 - able to be … in a few things 1 - unable to be at all …
Achievement and progress	I can achieve and progress in all aspects of my life	4 - can … in all aspects 3 - can … in many aspects 2 - can … in a few aspects 1 - cannot … in any aspects
Enjoyment and pleasure	I can have a lot of enjoyment and pleasure	4 - can have a lot 3 - can have quite a lot 2 - can have a little 1 - cannot have any

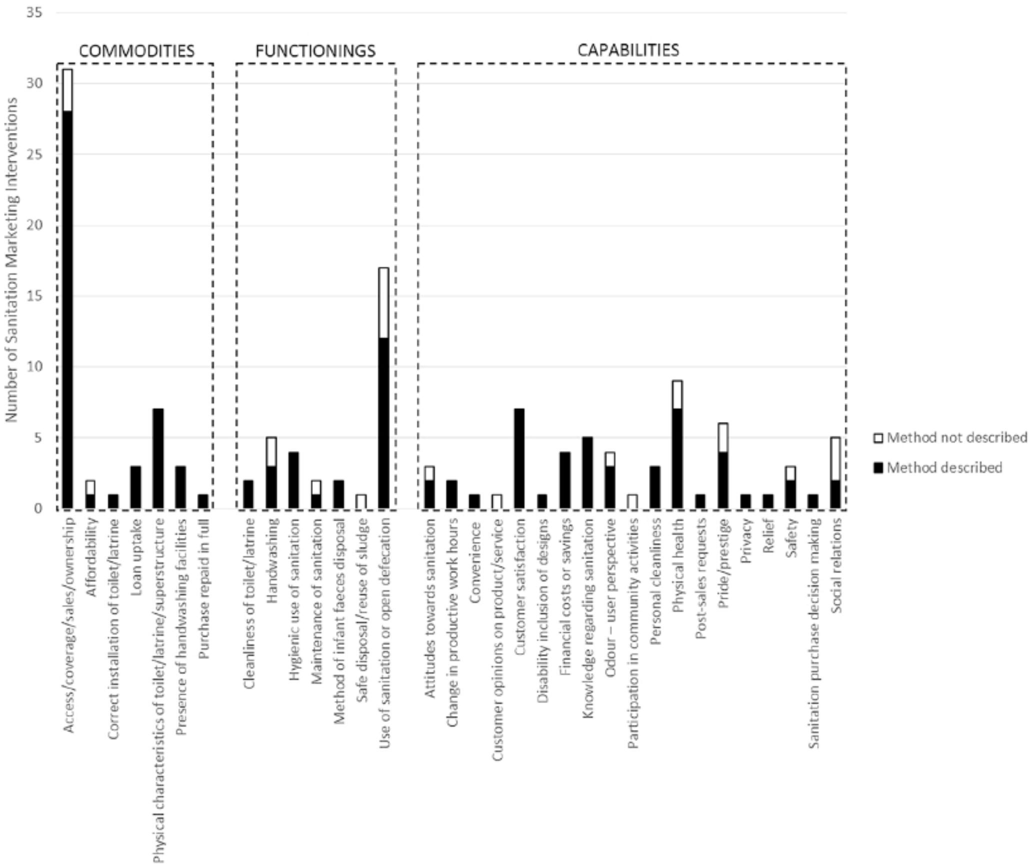
Some recent conceptual and empirical work has focused on what the decision rule for capability-based measures could be framed. The concept of “sufficient” capability has recently been developed (Mitchell et al., 2015). Whereas the goal of QALY-based CEA is to maximise health, this approach suggests that capability-based CEA might better focus on equity by using a decision rule that supports achievement of a normative minimum level of capability by all. The sufficient capability idea has been applied in a methodological case study, using “years of full capability equivalent” and “years of sufficient capability equivalent”, measured using ICECAP-A (Goranitis et al., 2017). A further study undertook a public deliberation process to establish define the "sufficient" state of capability well-being within the ICECAP-A descriptive system (Kinghorn, 2019). The participants settled on state 33333, that is, one step below the maximum for all attributes.

2.4.3. How has the capability approach been applied in sanitation?

There has been very little work applying the capability approach to sanitation. In fact, only one peer-reviewed study has done so, which was a systematic review applying the capability approach to assess sanitation marketing interventions (Barrington et al., 2017). The authors reviewed studies of 33 sanitation marketing interventions, noting that most had focused on toilets (commodities) but fewer on well-being outcomes (capabilities and functionings). Their main results graph is reproduced in Figure 4. The paper does not explain how capabilities and functionings were defined in the context of the research. From the figure, it appears that commodities were considered predominantly as characteristics of the toilet itself, and any handwashing facilities or loan characteristics. Functionings were considered as behaviours, such as toilet use, handwashing, child faeces disposal. Capabilities were considered as the well-being outcomes, e.g. privacy, safety, health.

The placing of capabilities as a consequence of functionings is unusual, since most capability literature frames things the other way around (Figure 3) (Robeyns, 2005). The authors justify this on the basis that using a toilet is a functioning, with other capabilities (e.g. health, privacy, safety) resulting from that functioning, which has a logic to it (Barrington et al., 2014). However, I prefer to follow the more standard approach (discussed below).

Figure 4: Concepts for which data was collected in 33 studies of sanitation marketing interventions (reproduced from Barrington et al., 2017)



A number of unpublished theses have addressed some aspect of sanitation and the capability approach. Three Masters theses linked to a team at the University of Basel focused on capabilities in relation to ecological sanitation (ecosan). One applied the capability approach in identifying drivers and barriers for the use of ecosan toilets in rural Nepal, using qualitative interviews with users and non-users (Messmer, 2011). The author found that the main reasons people gave for building ecosan toilets were related to food and related financial benefits, i.e. returns from increased production from application of urine as fertiliser. Another explored the capabilities of training participants in an ecosan programme in Pune, India, of limited relevance to this thesis (van Gelder, 2011). The third focused on ecosan in Tamil Nadu, India, but was not publicly available (Enssle, 2010).

Beyond ecosan, a PhD thesis assessed sanitation in peri-urban Chennai, India, using qualitative comparative case studies, by interviewing participants in 10 different “slum” areas (Immler, 2018). The author concluded that when participants practised open defecation, they were restricted in their freedom to be safe from emotional or physical harm, and experienced insecurity due to lack of privacy. Another Master's thesis, in Portuguese, explored how poor sanitation affects people's lives and freedoms, by interviewing 11 families in the city of Imperatriz, Brazil (Oliveira, 2014). The author found that personal security was particularly affected, since people are not able to control their own environment, as well as self-esteem.

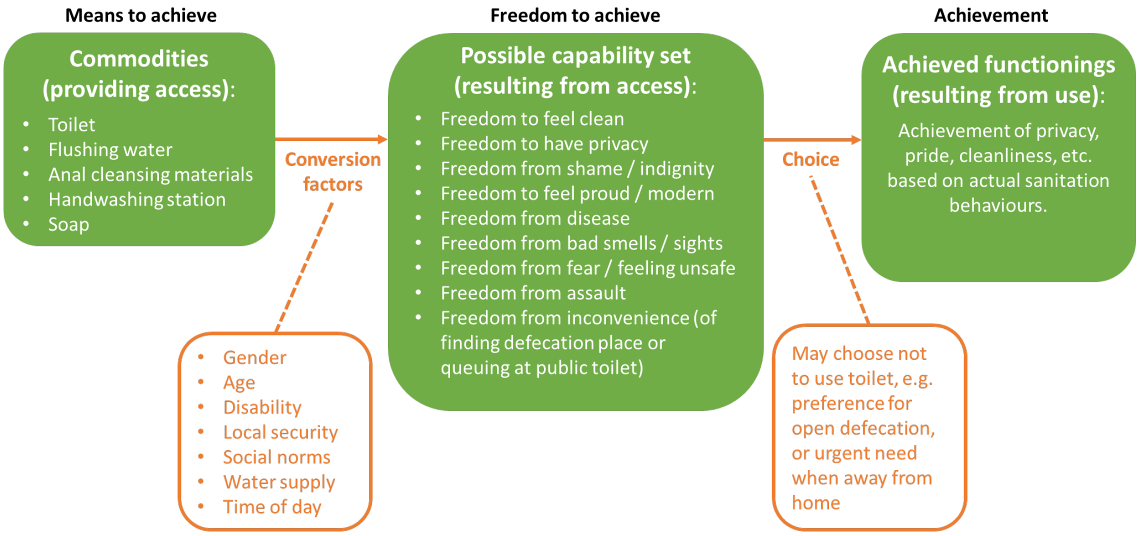
A number of unpublished conference proceedings also addressed sanitation and the capability approach. One study investigated "sanitation well-being" using interviews in urban Thailand (Rajbhandary et al., 2019). The authors identified anxiety, odour, privacy, safety, health and comfort as important themes relating sanitation to well-being. Another compared the "technologies for freedom" approach to community-led total sanitation in a theoretical exploration (González and Aristizábal, 2011). The authors suggest that "living in a healthy environment" can be considered a collective capability, since it cannot be achieved as an individual.

Summary of the state of the literature on sanitation and capabilities

I was able to draw only limited insight from the above theses and conference papers. All the empirical studies used qualitative methods, but were generally weak in explaining how the capability approach had influenced their methods and results. In terms of results, no new insights were added to what was already concluded from the literature on sanitation motives and drivers discussed in section 2.3 above. None aimed to develop any kind of capability-based measure.

Figure 5 presents a hypothetical set of sanitation commodities, capabilities and functionings, alongside conversion factors, adapting Robeyns' (2005) visualisation. The commodity (e.g. toilet) enables capabilities such as the freedom to have privacy. The individual chooses whether to use the toilet and actually achieving the privacy functioning in their practice of defecation, menstrual hygiene etc.

Figure 5: Sanitation commodities, capabilities and functionings (based on Robeyns, 2005)



Here are two examples of where an individual might choose not to use a toilet despite having access (cf. Sen's fasting/starving example). First, there is plenty of evidence of revealed preference for open defecation (Coffey et al., 2014) or reversion to open defecation (Crocker et al., 2017). Second, people may prefer to use buckets in their house at night-time, in settings perceived as unsafe at night. A qualitative study in Mozambique found this practice even in settings with high quality toilets, where women were afraid to walk to a shared toilet 5-20m from their house (Shiras et al., 2018). This represents an example of choosing not to act on the capabilities of privacy and avoiding embarrassment at using a bucket. In this case, those are overridden by a higher priority being placed on personal safety, a capability which has been temporarily lost at night-time. In a similar vein, preference for open defecation may change over time and with the state of the toilet – when a pit latrine is clean an individual may be happy to use it, but when it is smelly and full of flies, they may find it more pleasant to practice open defecation.

2.5. Conclusion

This chapter comprised an exploratory literature review. A formal systematic review with clearly defined search strategy is needed to identify whether and how QoL outcomes related to sanitation have been measured and valued in economic evaluation (Chapter 4).

This first part of this exploratory review has focused on economic evaluation as applied to health, economic evaluation as applied to sanitation, and the capability approach. Three key conclusions emerge. First, health economists have been keen to develop methods for the economic evaluation of interventions which influence outcomes beyond health as measured by QALYs, such as public health or social care interventions. However, most have continued to avoid welfarist CBA approaches, and explored other methods for extra-welfarist valuation based either on the "extending the QALY" or the capability approach.

Second, welfarist CBA appears to be the dominant economic evaluation method for assessing sanitation interventions. However, the user-reported benefits of sanitation related to quality of life, such as privacy, safety and dignity, have never yet been included in a CBA study. This is due to the absence of methods for measurement and valuation of these benefits, either in monetary or extra-welfarist terms. In short, no measure of the value of sanitation to users exists. Linked to this, sanitation has more often been considered through the lens of technologies and services, rather than a state of being.

Third, the capability approach has received little attention in the sanitation sector. Studies are limited to unpublished theses, conference papers, and a systematic review of sanitation marketing which used the capability approach as an organising framework. This comes despite the potential of the capability approach in assessing sanitation as a state of being. Much discourse on sanitation focuses on toilets and behaviours. Toilets can be characterised as commodities and behaviours as functionings. There has been less focus on the value of sanitation in terms of outcomes people experience, whether doings (ability to do) or beings (ability to be). For Sen, both beings and doings are part of the capability evaluative space. The capability approach is therefore attractive as a way to conceptualise QoL related to sanitation. Extensive existing work on capability measurement within health economics, across a diverse range of areas of health, shows that the capability approach can be flexible in application to economic evaluation.

Chapter 3: The setting in Maputo, Mozambique

3.1. Introduction

This chapter outlines the setting chosen to explore the research aim and objectives set out in Chapter 1, namely Maputo, Mozambique. The rationale for the selection of this setting was pragmatic – the MapSan trial offered an existing epidemiological study of a sanitation intervention in a low-income urban population with scope for supplementary activities to address my research objectives. Working within the structure of the MapSan trial provided important advantages of which three main ones are discussed here.

First, the trial was evaluating a specific sanitation intervention. The setting in urban informal settlements in a low-income country was relevant for the applicability of my methods and results in other similar settings. The intervention had already been delivered with fidelity, and quality of service had demonstrably improved. The hypothesis of whether quality of life had also improved could be tested, and an estimate of the effect combined with costs in an economic evaluation.

Second, the existence of intervention and control groups would be useful in informing both the qualitative and quantitative aspects of my research. For the qualitative research, it was useful to be discussing quality of life with people who experienced a substantially different quality of sanitation service. In addition, those in the intervention group could reflect on their present experience in relation to previous experience. For the quantitative research, the study design of the trial provided a basis for statistical inference about costs and outcomes. While it was several years too late for prospective data collection, a reasonable level of inference would still be possible even for data collected retrospectively within the two groups. Despite the limitations of this approach in terms of the risk of confounding, it was likely more informative than it would have been working in cross-section with people who had not necessarily received an intervention.

Third, there were practical benefits in working alongside a pre-existing structure. Strong collaborations already existed with Mozambican academic partners and a survey partner. One of my supervisors (Oliver Cumming) was the London School of Hygiene and Tropical Medicine (LSHTM) Principal Investigator for the trial, and grant-holder. The implementing agency, WSUP, and other local stakeholders were already working alongside the team.

There were limitations to this approach which I discuss in the concluding section of this thesis (Chapter 10). This chapter outlines, first, the context of Mozambique and Maputo City specifically. It then provides an overview of the urban sanitation policy environment in Mozambique and Maputo. It ends by describing the study design of the MapSan trial, the specific setting within Maputo in which it operated, and the evaluated intervention itself.

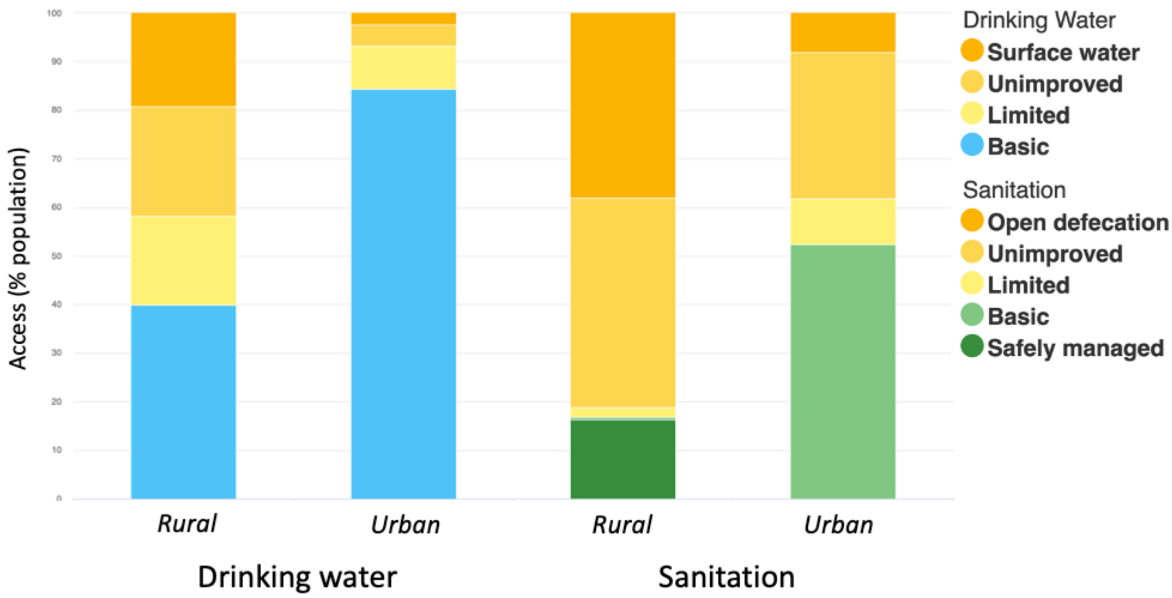
3.2. Mozambique and Maputo context

3.2.1. Mozambique

Mozambique is situated on the south-eastern coast of Africa, bordering six other countries. Portuguese is the official language, and also the most widely spoken at home (INE, 2019). Despite robust economic growth in recent decades, it remains one of the poorest countries in the world, ranking 180th out of 189 countries on the human development index (UNDP, 2019). Alongside high underemployment this suggests that growth, most recently driven by extractive industries, has not been inclusive (World Bank, 2017). There was sharp macroeconomic shock in 2016 when a secret public debt equivalent to 10% of GDP was uncovered, the consequences of which are still holding back the economy (Macuane et al., 2018).

While 63% of the population still lives in rural areas, the urban population is growing at 4.4% a year (United Nations, 2019). If this growth continues, the urban population will double every 16 years. Most of the urban growth has happened in the Greater Maputo Area which comprises a third of the national urban population. Poverty is reducing faster in cities than rural areas but 49% of urban households are in poverty by the World Bank headcount ratio definition (World Bank, 2017).

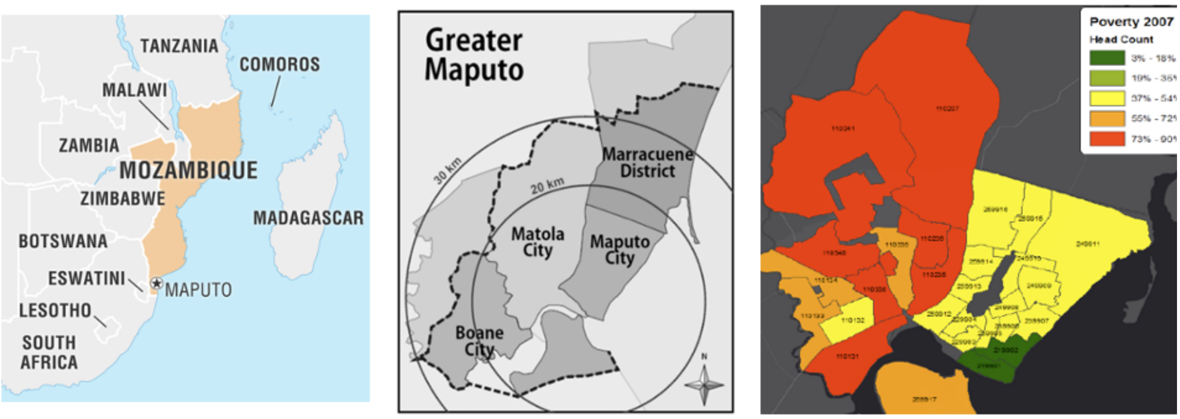
Figure 1: Access to water and sanitation services in urban and rural Mozambique (UNICEF & WHO, 2019)



Population growth will put increasing pressure on existing infrastructure and services. UNICEF and WHO (2019) report that in urban areas in 2017, 84% of people have basic drinking water, and 52% have basic sanitation (Figure 1). Basic urban sanitation access increased from 32% in 2000, but the improvement was unequal. The gap between richest and poorest quintiles increased by 30 percentage points (UNICEF & WHO, 2019). It is estimated that economic losses caused by poor sanitation in Mozambique are worth 1.2% of GDP annually, comprising the value of lost time, premature death, lost productivity and heath care costs (World Bank, 2012).

Figure 2: Maps of Maputo's location, administrative divisions and poverty headcount

A. Maputo within Mozambique | B. Greater Maputo region | C. Poverty headcount in Maputo and Matola (2007)



Encyclopædia Britannica (2019) | *JICA* (2014) | *World Bank* (2017)

3.2.2. Maputo

Maputo is Mozambique's capital, situated on the coast at Mozambique's southern tip, near the border with South Africa. The Greater Maputo Area is Mozambique's largest urban centre, with a population of 2.9 million according to the 2017 census (INE, 2019). Maputo City comprises 1.1 million people. The adjoining Matola City is now larger than Maputo City itself at 1.6 million. Matola doubled its population between 2007-2017 and is growing four times faster than Maputo (World Bank, 2017).

In 2007, poverty was above 40% in all areas outside the Maputo central business district (the green area in Figure 2, Panel C). The majority of people in Greater Maputo live in settlements with basic infrastructure, unpaved roads, and no drainage. Many people commute to Maputo City, by foot or by several journeys on packed minibuses (*chapas*). Access to piped water and electricity has increased substantially in recent decades (Figure 3), but is still lower the farther households are situated from Maputo city centre.

Figure 3: Access to infrastructure services by distance from Maputo city centre, 1997-2013 (World Bank, 2017)

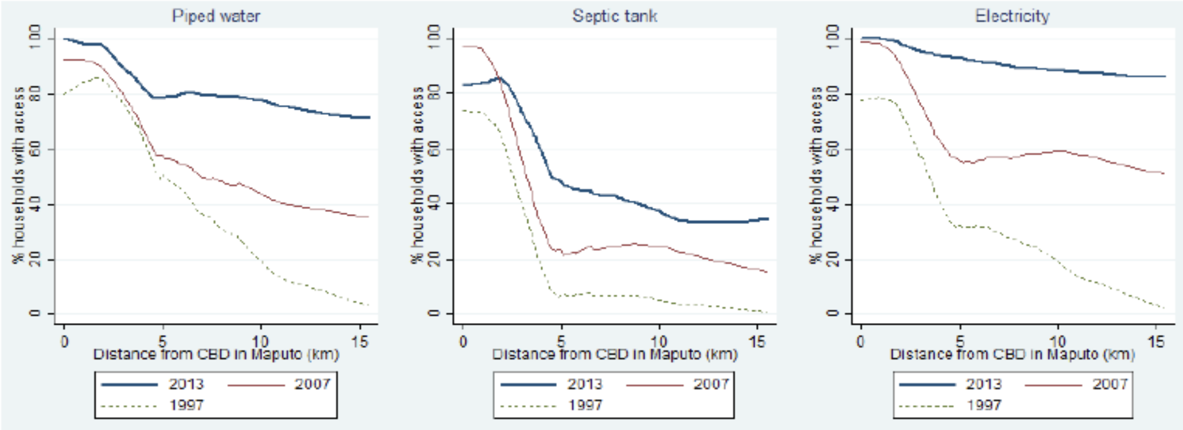


Figure 4: Access to sanitation in Maputo City by toilet type, based on Hawkins & Muximpua (2015)

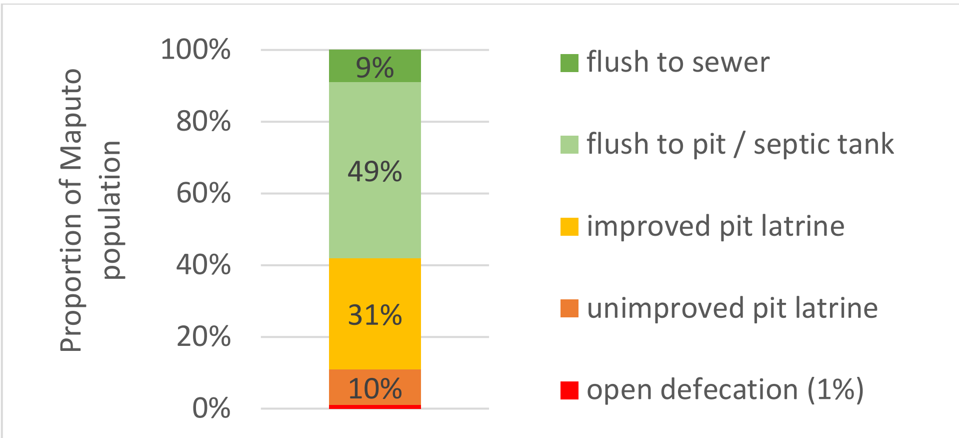
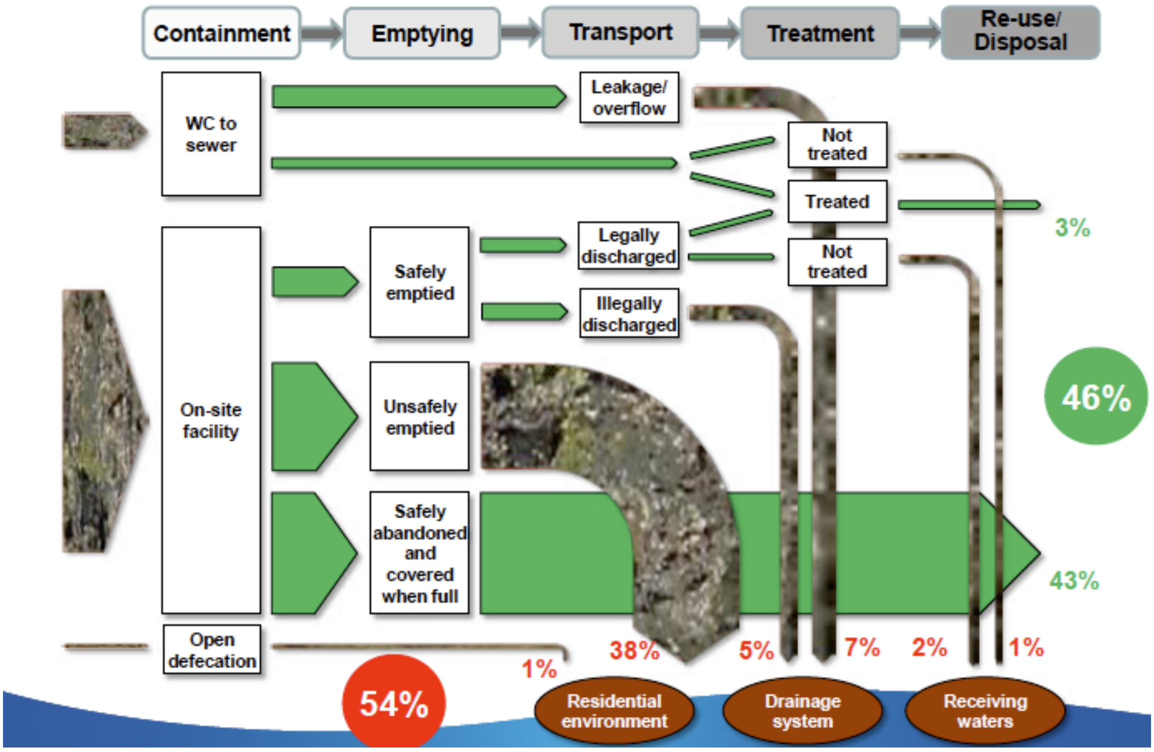


Figure 5: Faecal waste flows in Maputo (Hawkins and Muximpua, 2015)



Sanitation lags behind. Figure 4 shows that about 40% of the population of Maputo City uses pit latrines of varying quality, meaning toilets without a water seal (u-bend) which do not require flushing (Hawkins and Muximpua, 2015). Only 9% are connected to the sewer. The sewer network is far from being effective, with substantial leakage and overflow into the drainage system (World Bank, 2019a). When full latrine pits and septic tanks are emptied, this is often done unsafely, with the waste disposed into the environment or drainage system. Figure 5 depicts a fecal waste flow diagram, showing that only 46% of fecal waste generated in the city is safely managed (Hawkins and Muximpua, 2015). That proportion falls to 3% if an optimistic assumption is relaxed, namely that covering and abandoning full latrine pits comprises safe management in dense urban settings where flooding is common.

In 2014, the World Bank funded the development of a master plan for sanitation and drainage in Greater Maputo (AIAS, 2015). This was followed up by the 2019 approval of a

US$ 115 million World Bank project for improving urban sanitation across Mozambique, which should revolutionise the sector (World Bank, 2019a). It will fund sewerage investments in three cities including Maputo, as well as on-site sanitation and municipal service improvements. There is also a US$ 15 million technical assistance component which should increase strengthen municipal water and sanitation departments.

3.2.3. Urban sanitation policy environment

In urban areas, municipalities are responsible for ensuring provision of sanitation services. Under existing laws, this can be achieved either directly by the municipality or through an autonomous entity such as a utility (WSUP, 2019). In most cases, however, delivery has remained within municipal water and sanitation departments. The Maputo Municipal Council (CMM) recently approved a municipal sanitation policy (CMM, 2017). CMM has an annual infrastructure budget for water supply and sanitation, which was about US$ 3 million in 2017 (CMM, 2018). Since water supply is the responsibility of another agency, the budget is mostly allocated to stormwater drainage (55%) and sewerage (22%). Only 3% is dedicated to non-sewered sanitation, despite the fact that only 9% of the population is connected to the sewer, as is common in LMIC cities (Scott et al., 2019). Nonetheless, the Water Policy (Resolution 46/2007) and the National Strategy for Urban Water and Sanitation 2011-2025 promote a broad vision for urban sanitation, acknowledging the importance of on-site sanitation alongside sewerage.

Institutional responsibility for aspects of sanitation beyond service provision are as follows (WSUP, 2019). Policy development is the responsibility of the National Directorate of Water Supply and Sanitation (DNAAS) in the Ministry of Public Works, Housing and Water Resources. The ownership of public sanitation-related assets is split between the Water and Sanitation Infrastructure Management Agency (AIAS) and Municipalities. The intention is that AIAS would progressively transfer ownership to Municipalities as they increase their ability to manage the assets, though this is unclear. Regulation of sanitation is the responsibility of the Water Regulatory Authority (AURA), though practically there is lack of clarity in relevant laws about its remit in relation to Municipalities (WSUP, 2020).

From the perspective of economic evaluation of sanitation options in most cities in Mozambique, the relevant decision-maker allocating a fixed sanitation budget is the municipality. This is also the case in Maputo specifically.

3.3. The Maputo Sanitation (MapSan) Trial

Research undertaken for this thesis took place alongside the Maputo Sanitation (MapSan) trial (clinicaltrials.gov registration: NCT02362932). It was the first controlled trial of urban non-sewered sanitation. It was also the first sanitation trial to use a measure of enteric infections as the primary outcome, rather than a measure of disease.

3.3.1. MapSan trial study design

The MapSan trial evaluated the impact of a shared urban sanitation intervention on children's enteric infections (Brown et al., 2015). Specifically, it evaluated an intervention (described below) implemented in 2015-16 by Water and Sanitation for the Urban Poor (WSUP), an international non-government organisation (NGO). The intervention was funded as part of a broader programme funded by the World Bank.

The study used a Controlled, Before-and-After design. As each intervention compound was identified, a control compound was matched with it, based on cluster size and intervention siting criteria (Brown et al., 2015). Such a design cannot eliminate the risk of residual confounding due to unmeasured or unknown confounders. Covariate balance was achieved at baseline (Knee et al., 2018), at which time the study population at baseline comprised 447 under-5 children in an intervention group and 536 in a control group. The MapSan trial concluded that the intervention did not reduce the overall prevalence of enteric infection and diarrhoea among enrolled children (Knee et al., 2020).

Figure 6: Location of MapSan intervention and control compounds within Maputo City

3.3.2. Setting

The MapSan trial setting is low-income neighbourhoods in the Nhlamankulu district of southern Maputo. In these neighbourhoods, housing is diverse but the poorest people live in informally-walled ‘compounds’, where many families share the same toilet and courtyard space (Brown et al., 2015). The location of MapSan compounds within Maputo City is shown in Figure 6, comprising a relatively a small area of about 10km^2 within the Nhlamankulu district.

Low-quality self-built pit latrines (PLs) are common in these compounds. These vary widely in their quality, but many comprise an informally-walled area of a compound rather than a building, providing little privacy (Capone et al., 2020). These are typically unlined pits with squatting slabs made of wood, tyres or concrete, and no water seal (u-bend) providing a barrier to smells and flies (Mattson, 2016). Few have roofs, and the walls are made with sections of scrap corrugated iron or plastic sheeting, with makeshift fabric doors – see photos in Table 1. When such pits fill, users sometimes cover them and dig a new one, or empty them if there is no space (Capone et al., 2020). Drainage is poor in these areas, and flooding in seasonal rains risks pits overflowing into compounds and streets.

Table 1: Photos of pit latrines on control compounds

1. Pit latrine with tyre/wood for squatting	2. Pit latrine with concrete slab
	
3. Fabric door providing limited privacy	4. No door and adjacent greywater pit
	

3.3.3. Intervention

The intervention provided a subsidised block-built pour-flush toilet (i.e. with a water seal) discharging to a septic tank with soakaway. All toilets were designed to be shared by multiple households, given the arrangements on compounds. There were two design types depending on user numbers. The first was a shared toilet (ST) with one stance (cubicle) to be used by around 15 people, at 85% subsidy. The second was a Community Sanitation Block (CSB) with two stances, to be used by a minimum of 21 people, at 90% subsidy. Both STs and CSBs are robustly built, with doors lockable from the inside. The engineering design of a CSB is shown in Figure 7. The ST was a simpler design, without a rooftop water tank or laundry station, and only one stance. All CSBs included handwashing basins but only some STs did so, depending on design.

By contrast to PLs, the ST and CSB options are built of cement blocks, have a metal roof, concrete floor, and a door which locks from the inside. Both designs include a handwashing basin. The 'pour-flush to septic tank' interface prevents smells. See photos in Table 2.

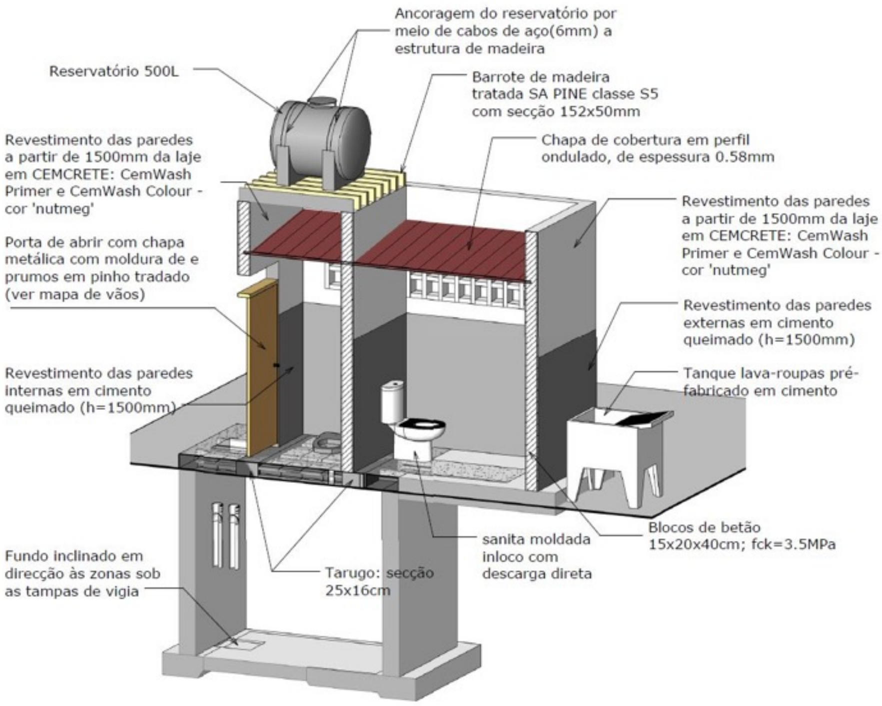
Compounds were eligible for the intervention if households were sharing a low-quality pit latrine and willing to pay the 10-15% capital contribution – approximately US$ 120 for CSB (2015 prices) and US$ 80 for ST (Mattson, 2016). Note that these toilets are still shared. Much sanitation literature, and the sustainable development goal targets, focus on the user interface and waste treatment. However, many aspects of the overall toilet design contribute to privacy and safety without any likely infectious disease consequences, such as the solid walls and lockable door. The water storage in the CSB design also facilitates handwashing and menstrual hygiene. In what follows, I use "toilet" to refer to the whole infrastructure, not only the user interface.

The NGO contracted eight community-based organisations to engage potential users, select sites and collect financial contributions. Nine local construction firms were contracted to prepare sites and build toilets. The World Bank and CMM provided oversight. During 2015-16, the intervention was delivered to 450 compounds. CSBs were installed in 50 compounds and STs in 400, benefitting approximately 7,000 people overall. Control households did not receive an intervention, but were free to upgrade their sanitation option autonomously if they desired to do so.

Table 2: Photos of toilets on intervention compounds

Exterior	
1. Shared toilet (ST)	2. Community sanitation block (CSB)
	
Interior (varied between CSB / ST depending on design)	
3. Squat pan	4. Seat pan

Figure 7: Engineering design of community sanitation block with Portuguese annotations (source: WSUP)



Chapter 4: Research Paper 1 - Systematic review

Chapter introduction

As identified in Chapter 2, there has been no in-depth review of economic evaluations of sanitation interventions, which was the primary motivation for undertaking the study reported in this paper-style chapter. For two reasons, interventions to improve drinking water were incorporated in the protocol, including both water supply and household water treatment. First, drinking water and sanitation interventions are sometimes delivered together, and I was aware of a number of economic evaluations which studied combined interventions. Second, drinking water presents similar methodological challenges to sanitation, as a "public health engineering" sector in which services are delivered based on long-lived assets which contribute benefits within and beyond health.

The study reported in this chapter (Research Paper 1) presents a systematic review of the cost-benefit and cost-effectiveness of drinking water and sanitation Interventions in low- and middle-income countries. It aimed to fulfil objective 1 of my thesis set out in section 1.3: to assess the extent to which quality of life outcomes have been measured and valued in economic evaluations of sanitation interventions. I am proposing to submit the paper to *BMJ Global Health* (published by BMJ Journals) and, accordingly, it follows the standard structure for biomedical journals: introduction, methods, results and discussion. The two main results tables and Supplementary Materials are in Appendix A at the end of this thesis.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1701671	**Title**	Mr
First Name(s)	Ian		
Surname/Family Name	Ross		
Thesis Title	Measuring and valuing quality of life in the economic evaluation of sanitation interventions		
Primary Supervisor	Giulia Greco		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMJ Global Health
Please list the paper's authors in the intended authorship order:	Ian Ross, Louise Watson, Giulia Greco, Barbara Evans, Marc Jeuland, Oliver Cumming
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceptualised the study and developed the protocol. I coordinated inputs into the protocol and manuscript from my supervisory team and external co-authors. I designed and implemented the search strategy. I led the screening, data extraction and quality scoring processes, coordinating the inputs of the second reviewer into these. I conducted the analysis and write-up, revising drafts on the basis of co-author feedback.

SECTION E

Student Signature	
Date	24/10/20

Supervisor Signature	
Date	24/10/20

Cost-benefit and Cost-effectiveness of Drinking Water and Sanitation Interventions in Low- and Middle-Income Countries – a Systematic Review

Authors:

Ian Ross[1], Louise Watson[2], Giulia Greco[1], Barbara Evans[3], Marc Jeuland[4], Oliver Cumming[1]

Affiliations:

[1]London School of Hygiene and Tropical Medicine, UK;

[2]University of Waterloo, Canada;

[3]University of Leeds, UK;

[4] Duke University, USA;

Corresponding author:

Ian Ross

LSHTM, 15-17 Tavistock Place, London, WC1H 9SH

Email: ian.ross@lshtm.ac.uk

Phone: 020 7636 8636

Abstract

Background: Poor water supply and sanitation contribute to a substantial disease burden in low- and middle-income countries (LMICs) as well as negative economic and social impacts. The water and sanitation sector receives significant investment, but economic evaluations to guide decisions remain scarce. We aim to synthesise evidence on the cost-benefit and cost-effectiveness of drinking water and sanitation interventions in LMICs.

Methods: We searched peer-reviewed and grey literature published since 1980 using databases, websites, and reference lists of included studies. We included studies only if they reported "full" economic evaluations, which measure both costs and outcomes in a cost-effectiveness or cost-benefit analysis. Interventions included water supply, water quality, or sanitation. We scored study quality based on the CHEERS checklist.

Results: We identified 53 publications, of which 19 were scored as being "high" or "very high" quality. Only 20 studies evaluated empirical interventions, defined as economic evaluations of interventions which were actually implemented. For interventions providing communal boreholes in rural areas, benefit-cost ratios (BCR) between 2 and 5 appear realistic. For rural sanitation campaigns, BCRs between 1 and 3 appear realistic. However, uncertainty and heterogeneity mean that BCRs below 1 remain plausible if interventions are poorly targeted at areas with limited potential for uptake. Too few studies of sufficient quality were identified to draw similar conclusions about other intervention types or cost-effectiveness studies. Saved time was the benefit most commonly measured and valued in cost-benefit studies, followed by direct health benefits. No studies directly valued quality of life outcomes, such as safety, privacy and dignity.

Conclusion: Drinking water and sanitation interventions do not always deliver net benefits. Economic evaluation is important to inform more efficient resource allocation, but the small number of high-quality studies impedes efficient decision-making. In particular, cost estimation, measurement of non-health benefits, and methodological transparency could be improved. Funders of impact evaluations may consider requiring high-quality accompanying economic evaluations to inform future decision-making. Future studies could be better tailored to decision problems regularly challenging WASH professionals in LMICs, for example by assessing multiple alternative interventions.

Introduction

In 2017, 580 million people globally used unimproved water sources for drinking, and two billion were without basic sanitation (UNICEF & WHO, 2019). Of the 1.4 million deaths from diarrhoeal disease in 2016, 485,000 were attributable to inadequate water supply and 432,000 to inadequate sanitation (Prüss-Ustün et al., 2019). Inadequate water, sanitation and hygiene (WASH) services are also associated with poor mental well-being (Sclar et al., 2018). Annual economic losses from inadequate WASH are estimated to reach 6% of gross domestic product (GDP) in many countries (Hutton and Chase, 2016). This economic burden continues to rise in absolute terms in Sub-Saharan Africa and parts of South Asia (Jeuland et al., 2013)

There is increasing evidence concerning interventions' relative effectiveness in encouraging uptake of behaviours and technologies (Garn et al., 2017) and in preventing diarrhoea and other diseases (Wolf et al., 2018). To maximise the benefits achieved with constrained budgets, however, decision-makers must consider the relative costs of interventions, as well as their relative effectiveness. "Full" economic evaluations aim to promote more efficient resource allocation by comparing the relative costs of interventions to their relative benefits (Drummond et al., 2015). Full economic evaluation methods include cost-effectiveness analysis (CEA) and cost-benefit analysis (CBA), defined in Table 1.

Appropriate methods for economic evaluation of WASH interventions have long been debated (Briscoe, 1984; Churchill et al., 1987; Cvjetanovic and Grab, 1976). There has been no systematic review of economic evaluations for drinking water and sanitation, despite around $100 billion being invested annually in the WASH sector in LMICs (WHO, 2017). A recent review (Hutton and Chase, 2016) was not systematic, did not assess study quality, and did not compare the efficiency of specific intervention types. A more in-depth assessment and synthesis of existing economic evaluations of WASH interventions would help inform efficient resource allocation decisions and identify priorities for future research. A systematic review of economic evaluations provides a structure for such syntheses (Gomersall et al., 2015; van Mastrigt et al., 2016). In this article, we aim to synthesise evidence on the cost-benefit and cost-effectiveness of drinking water and sanitation interventions in LMICs.

Table 1: Definitions of key terms

Term	Definition
Cost-benefit analysis (CBA)	A method for comparing costs and outcomes, which values all outcomes in monetary terms and aggregates into total benefits. Results of a CBA are commonly expressed as a benefit cost ratio (BCR), net present value, or internal rate of return. The decision rule for CBA is that, if BCR > 1, the intervention is economically beneficial. BCRs for different interventions in the same setting can be compared to assess an efficient use of resources from the perspective evaluated.
Cost-effectiveness analysis (CEA)	A method for comparing costs and health or wellbeing outcomes, which values outcomes in 'natural' units (e.g. diarrhoea case, death) or composite units (e.g. disability-adjusted life years [DALYs]). Results of a CEA are commonly expressed as an incremental cost-effectiveness ratio (ICER). When one intervention has both lower costs and better outcomes than another, it is said to be "dominant" and costs and effects may be presented separately. The decision rule for CEA is that, if the ICER < a threshold, the intervention is considered an efficient use of resources; however, there is much debate and uncertainty about appropriate thresholds in a given context, and they are generally only generated for composite outcome metrics such as DALYs.
Technology	The infrastructure or product resulting from an intervention. Examples include a pit latrine, a borehole with handpump, or a bottle of chlorine solution. Consistent use of a technology is a behaviour.
Intervention	The mechanism for encouraging behaviour change or the direct provision of assets or services. Examples include "infrastructure delivery" interventions, which construct assets directly, and "campaign" interventions which promote uptake of behaviours and/or assets.
Hypothetical vs. empirical intervention	Studies of hypothetical interventions evaluate interventions which haven't actually been implemented. Such studies normally construct scenarios wholly or partly from secondary data, sometimes supplemented by some primary data from the hypothesised setting. Studies of empirical interventions collect primary data alongside and/or following actual implementation of an intervention, often supplemented by secondary data.
Specific vs. generic setting	Hypothetical interventions can be modelled for a specific country and setting (e.g. rural Ethiopia) or a generic non-specific setting. This can be at the global level, or an assumed typical setting. All empirical interventions (see above) are by definition in a specific setting.
Costing perspective	The perspective of an economic evaluation determines whose costs are included. Taking a societal perspective means including costs borne by all stakeholders (e.g. recurrent costs borne by households). Taking a provider perspective, uncommon for WASH economic evaluations, means including only costs borne by the service provider.
Full vs. partial costing	We define a "full" costing as including all important cost categories typical in WASH interventions, such as: capital, programme delivery, and recurrent costs. We define "partial costing" as one which omits one of more of these components, and "unclear" as not explaining which costs are included in an estimate.

Methods

The protocol was pre-registered with the International Prospective Register of Systematic Reviews (PROSPERO) - CRD42020165669. We report methods and findings (Figure 1) in accordance with PRISMA (Moher et al., 2009) and best practice for systematic review of economic evaluation evidence (Shemilt et al., 2019; van Mastrigt et al., 2016).

Eligibility criteria

We sought to identify full economic evaluations (Table 1) in any LMIC population which evaluated interventions in water supply and distribution, water quality, and sanitation. We defined LMICs according to the 2019/20 World Bank classification (Supplementary Material B). We included interventions targeting individuals, households and communities, but not those targeting institutions such as schools and healthcare facilities. Outcomes compared include any measures that are the output of full economic evaluation, such as benefit-cost ratio or cost per disability-adjusted life year (DALY) averted.

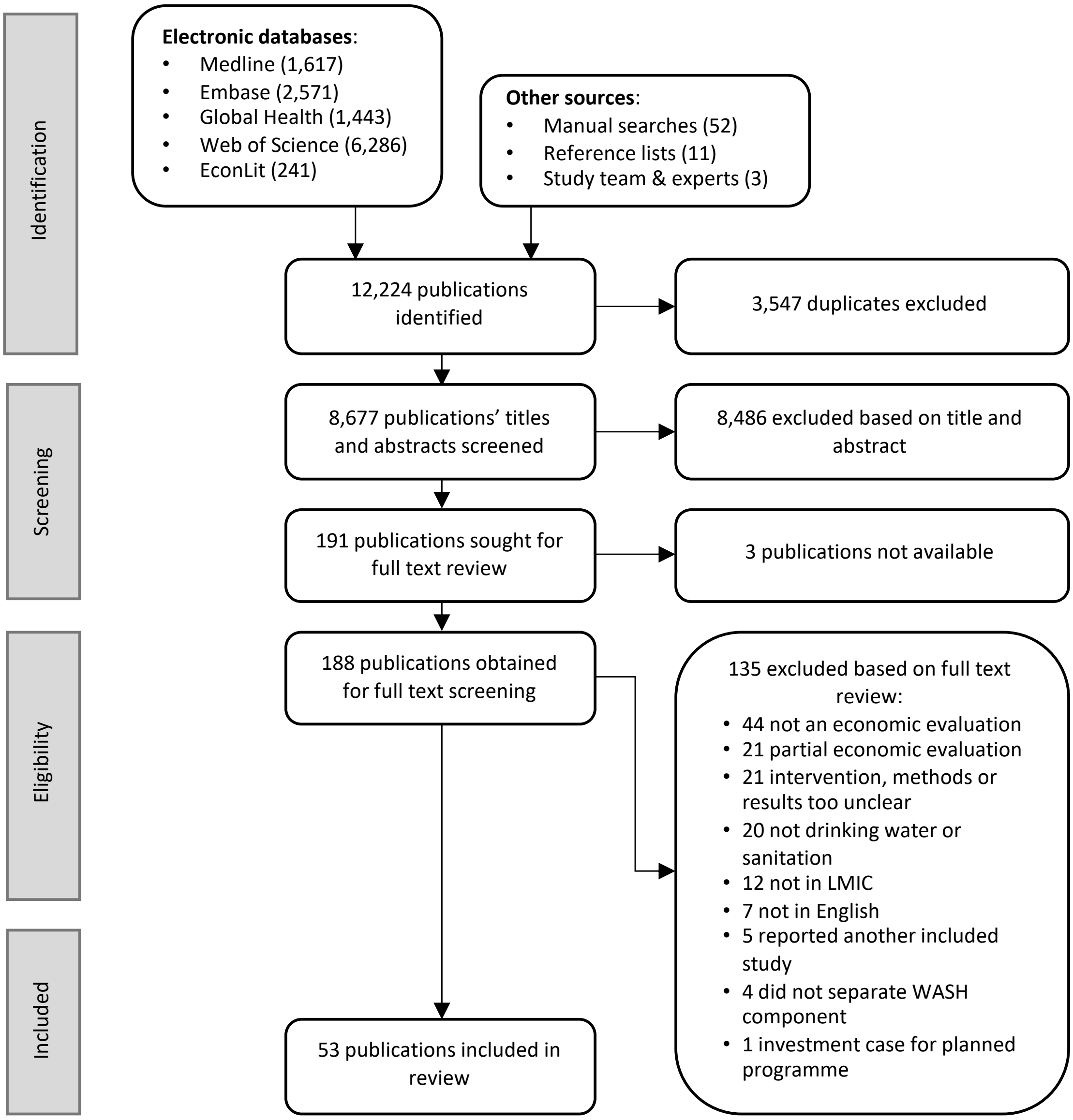
Information sources and searches

We searched titles and abstracts of peer-reviewed and grey literature for publications since January 1980, combining terms for (A) CBA or CEA, with terms for (B) water supply and distribution, or water quality and treatment, or sanitation (search strategy in Supplementary Material A). On April 20, 2020, we searched five bibliographic databases: Medline, Embase, Global Health, EconLit and Web of Science. We also searched 17 additional websites and databases, including the National Bureau of Economic Research and the Copenhagen Consensus Centre (full list in Supplementary Material A). Reference lists of included publications and any relevant reviews were also screened. Finally, we contacted 53 individuals who were either experts in the field or based at institutions active in the field, and asked them to identify published or ongoing studies and to suggest other experts we should approach. The email only requested a response if the expert had a suggestion for a study to include. Nine experts responded with a proposal, but all suggested studies were either already included or did not meet the inclusion criteria.

Study selection

Titles and abstracts were uploaded into Mendeley software (Elsevier, 2019) and independently screened for eligibility by two reviewers (IR and LW). The few discrepancies between reviewers were resolved through discussion. If citations appeared eligible or had the possibility of being eligible, they continued to full text review (Figure 1). We tabulate studies identified as ineligible upon review of the full text by reason for exclusion (Supplementary Material D).

Figure 1: PRISMA flow diagram for selection of studies



Data extraction

We extracted data items related to interventions, methods used, and results, based on guidance (Gomersall et al., 2015; van Mastrigt et al., 2016) and previous systematic reviews of economic evaluations for health interventions in low and middle-income countries (LMICs) (Mangham-Jefferies et al., 2014; Remme et al., 2014; Turner et al., 2019b). Two reviewers (IR and LW) piloted data extraction for three of the retained studies using a pre-agreed spreadsheet. One reviewer (IR) extracted data from included studies and a second reviewer (LW) assessed the accuracy of extracted data for 20% of studies. Only three discrepancies were identified, and resolved through discussion.

Quality assessment

We assessed quality using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS), which provides a set of 24 criteria (Husereau et al., 2013). While CHEERS was designed to promote transparent reporting, we use it as a proxy for overall study quality, as have many others (Hope et al., 2017; Mangham-Jefferies et al., 2014; van Mastrigt et al., 2016). We followed the CHEERS scoring strategy used in a previous systematic review of health interventions in LMICs (Mangham-Jefferies et al., 2014). Studies were awarded 1 point for each CHEERS item that was fully met, 0.5 for each partially met, and 0 when not met or insufficient information was reported (supplementary material C). We calculated a percentage score, giving all criteria equal weight. If an item was not applicable to a study (e.g. the "preference-based measures" criterion for studies which did not use them), that item was excluded from the scoring denominator for that study. Studies scoring 80% or more were categorised as "very high" quality, those scoring 65-79% as "high", 50-64% as "medium" and less than 50% as "low". Cut-offs were based on a previous systematic review (Mangham-Jefferies et al., 2014) and the histogram of scores. Two reviewers (IR and LW) piloted scoring for three studies. Thereafter, one reviewer (IR) scored all studies and a second reviewer (LW) reviewed scoring for 20% of studies. Two item scores were queried in two different studies, and resolved through discussion.

Synthesis and reporting

We present the extracted data in two types of tables, the first showing characteristics of included studies, and the second summarising their results. We provide separate tables for cost-effectiveness and cost-benefit studies, and separate tables for studies of hypothetical and empirical interventions within those. Throughout, we distinguish between: (i) interventions and technologies, (ii) hypothetical and empirical interventions, (iii) specific and generic settings, (iv) full and partial costing (Table 1). To aid comparison, we converted all cost-effectiveness estimates to constant 2019 United States dollars ($) using World Bank (2020) data on GDP deflators and exchange rates (supplementary material G). We conducted a narrative synthesis to qualitatively assess the overall strength of the evidence, and to discuss the strengths and weaknesses of methods used. Key terms are defined in Table 1, and we explain different types of uncertainty and heterogeneity in Supplementary Material H. Our results are arranged around: (i) study methods and quality; (ii) types of benefits measured and valued; (ii) comparison of study results. We only directly compare results for intervention types evaluated by at least three studies which were fully costed (Table 1) and rated "high" quality or above, as a normative threshold. For these comparisons, we presents results in forest plot style without meta-analysis.

Results

Identification, screening, eligibility, and inclusion

Electronic databases yielded 12,158 results with a further 66 identified from manual website searches, reference lists and expert consultation. After removing duplicates, titles and abstracts of 8,677 unique publications were screened, and 191 full texts reviewed (Figure 1).

At full-text review, 135 publications were excluded (Figure 1, supplementary material D). Forty-four publications were judged not to be economic evaluations. A further 21 were partial economic evaluations, reporting only costs (9) or only benefits (12). Twenty-one studies appeared to conduct full economic evaluations but were excluded because insufficient information was included to enable assessment of one or more of the intervention, the methods, or the results. Twenty studies reported on interventions that were not drinking water or sanitation, such as water resources management or wastewater reuse for irrigation purposes. A further 50 publications were excluded for other specified reasons (Figure 1, supplementary material D).

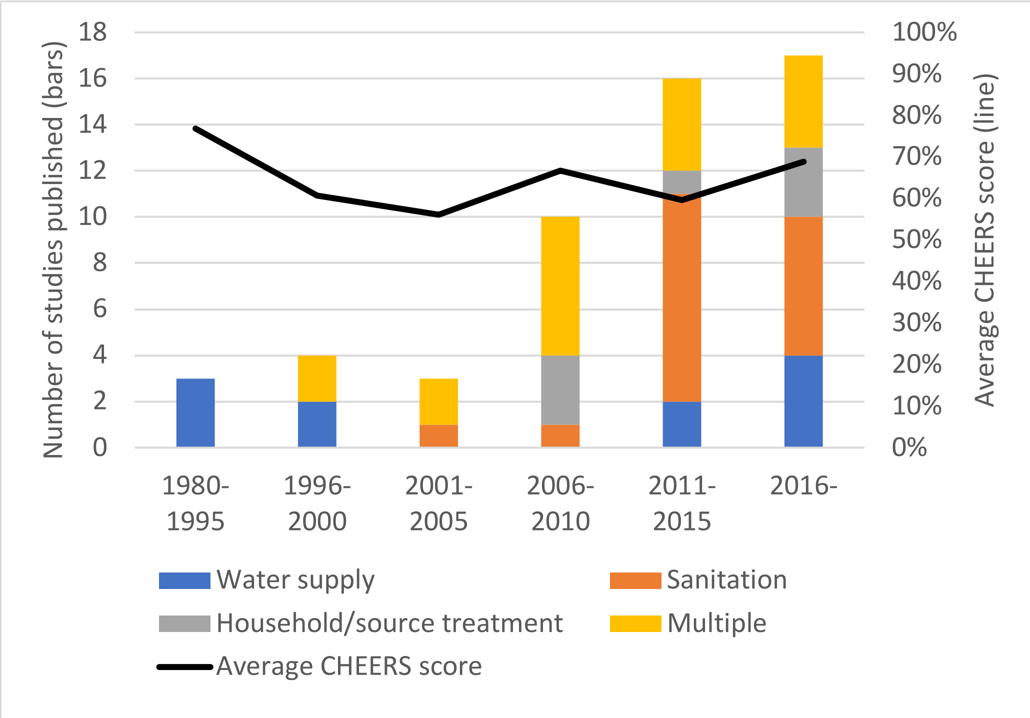
Description of included studies

Study characteristics and results are provided across six tables in Supplementary Material A:

- Table 2: Characteristics of included studies – empirical interventions
- Table 3: Characteristics of included studies – hypothetical interventions
- Table 4: Study results, empirical cost-benefit analyses
- Table 5: Study results, hypothetical cost-benefit analyses
- Table 6: Study results, empirical cost-effectiveness analyses
- Table 7: Study results, hypothetical cost-effectiveness analyses

Of the 53 publications included in the final review (Tables 2-3), 30 were peer-reviewed journal articles and 3 were academic masters theses. The remaining publications were reports commissioned by international agencies or working papers. In nine publications, the economic evaluation was not the primary focus of the study, but was included in a short section accompanying an impact evaluation. The number of full economic evaluations published increased substantially in the mid-2000s, especially for sanitation after 2010 (Figure 2). We present studies' characteristics (Tables 2-3) according to definitions set out previously (Table 1).

Figure 2: Number and quality of included studies (1980-2020)



nb. Only the 53 included studies are presented in this figure. Excluded studies, with reasons for exclusion, are listed in Supplementary Material D.

Studies most commonly evaluated an intervention versus a no intervention comparator. However, 20 studies evaluated two or more interventions against a comparator. Of the 53 studies, 34 evaluated an intervention for improving sanitation, 30 water supply, 12 household water treatment (HWT), and two source water treatment.

Technologies evaluated were diverse. Amongst studies which evaluated interventions in a specific setting (Table 1), the most common technologies evaluated for water supply were boreholes equipped with handpumps (n=8), followed by public taps with unclear source (n=4), and piped connections with unclear source (n=1). For HWT, common technologies were household biosand or ceramic filters (n=4) followed by household chlorination (n=3). For sanitation, the most common technologies were pit latrines (n=12), followed by flush toilets to pits or septic tanks (n=13) and sewerage with household connections (n=5).

Study design and quality

Study design

Of the 53 studies, 31 used CBA only, 15 CEA only, and seven used both. Thirty-three evaluated hypothetical interventions (Table 1), assumed to happen in a specific or generic setting. Some of these modelled scale-up to given coverage at the global level (Hutton et al., 2007a). Only 20 studies reported evaluations of empirical interventions which were actually implemented. Seven of these used outcome data collected under an impact evaluation of the same intervention, of which three were HWT studies (Burt et al., 2018; Cha et al., 2020, 2018; Dupas et al., 2020; Meddings et al., 2004; Rogers et al., 2019; Shrestha et al., 2006). The other 13 used primary cost data combined with effectiveness assumptions from secondary data.

Perspective and comparators

Almost all studies claimed to undertake economic costing from a societal perspective, including all costs regardless of the bearer (Table 1). Nine studies, which were mostly published earlier, applied a health provider perspective; most of these evaluated smaller-scale or limited technology interventions that health providers might feasibly deliver (e.g. chlorine distribution). Studies generally estimated costs and outcomes of interventions as incremental over a "without" intervention scenario (Tables 4-7 in Supplementary Material A). Only ten studies evaluated costs and outcomes as incremental over another feasible intervention (Cha et al., 2020; Chuan et al., 2012; Heng et al., 2012; Jeuland et al., 2009; Nguyen et al., 2012; Rodriguez et al., 2013, 2011; Rogers et al., 2019; Shrestha et al., 2006; Winara et al., 2011). Six of the latter were part of the same multi-country study in East Asia (Hutton et al., 2014).

Measures of comparison

A BCR was reported by 26 studies, making it the most common efficiency measure used (Tables 4-7 in Supplementary Material A). Other CBAs (n=5) reported net present value only. The majority of CEA studies (n=13) report cost per DALY averted, followed by cost per death or case averted (n=8). One study reported an intervention-specific measure: “cost per additional child recovered” (Rogers et al., 2019). Of the 31 CBA studies, 30 reported positive net benefits (BCR>1) under at least one scenario, with only one study not doing so (Sklar, 2017).

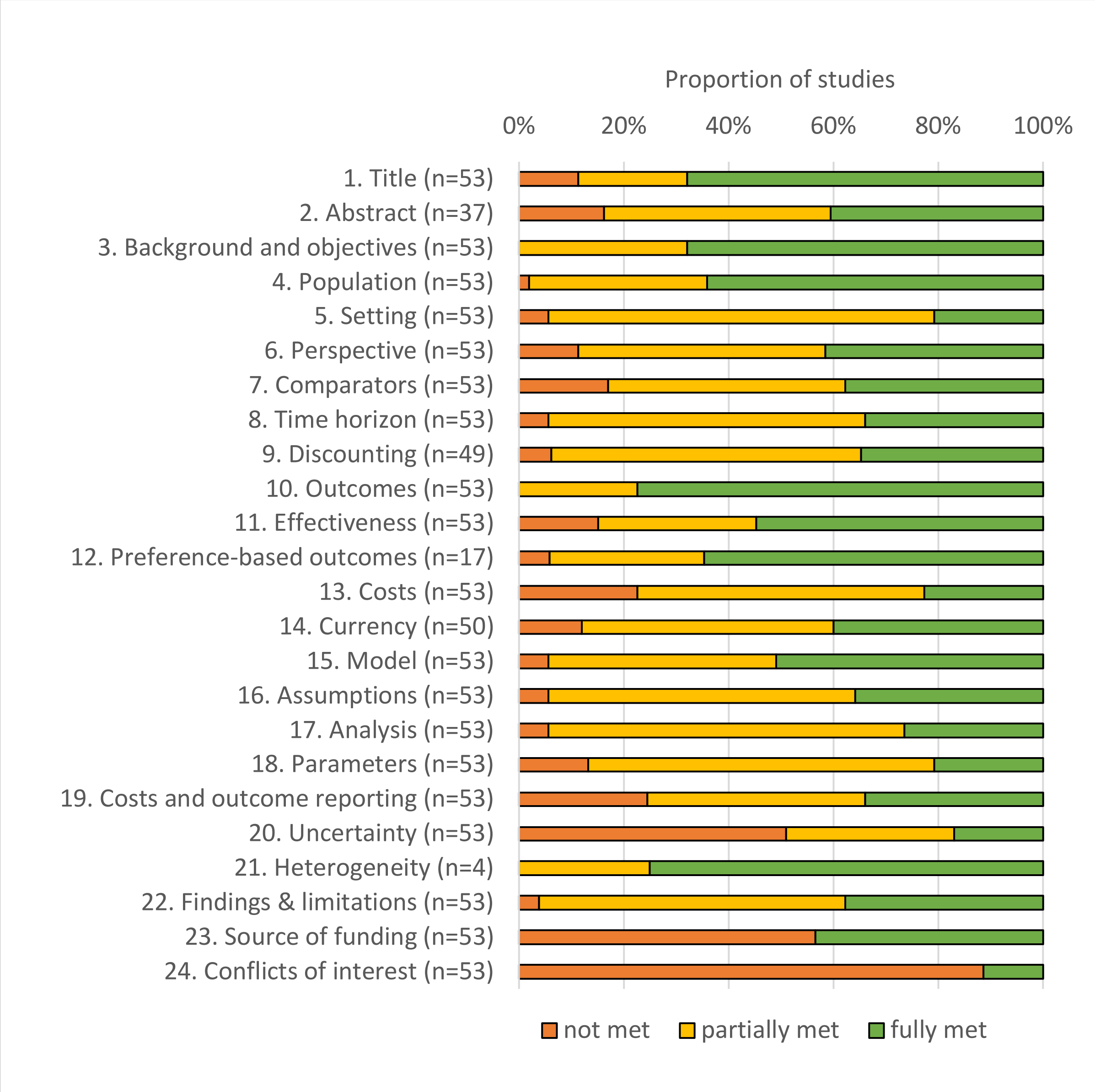
Sensitivity analysis

Thirty-three studies undertook deterministic sensitivity analysis (DSA) only and 12 undertook no sensitivity analysis at all. Only eight studies undertook probabilistic sensitivity analysis (PSA). Relatively few studies explored heterogeneity between population sub-groups (supplementary material H), usually by wealth quintile (Burt et al., 2018; Hutton et al., 2020). Some studies used Monte Carlo simulation to explore heterogeneity between locations and structural assumptions in the model (Supplementary Material H)

CHEERS scoring

The median score using CHEERS was 61% (range 32%-93%, Supplementary Material E). Eight studies (15%) were scored as very high quality, 11 (21%) as high, 26 (43%) as medium, and eight (15%) as low, based on the normative cut-offs we applied. Amongst the methodological criteria (#6-#21 in Figure 3), studies tended to meet the criteria for describing outcomes used (#10) and sources of effectiveness data (#11). They were least likely to meet the criteria for the characterisation of uncertainty (#20) and reporting of costing (#13). Many used aggregated secondary cost data (e.g. single figure for total expenditure on a programme) without discussing which costs might be included or excluded. Reporting the values, ranges and references for input parameters (#18) was also generally poor. On average, the quality of published economic evaluations has not improved over time (Figure 2).

Figure 3: Scoring of studies against CHEERS criteria



Types of benefits measured and valued

The 22 CEAs all used a health outcome as the denominator of the incremental cost-effectiveness ratio (e.g. cases averted, DALYs averted). The 38 CBAs measured and valued a wide range of outcomes to include in their aggregate measure of monetary benefits. We present two tables focusing on the CBA studies in Supplementary Material H. The first lists and explains the types of benefits commonly quantified, and the methods used for valuation of those benefits. The second presents the outcomes valued in each of the CBA studies, the percentage of benefits each contributes to the total for that study, and the methods used for valuing avoided mortality and time savings.

Of the 38 CBAs, 34 measured and valued saved travel time from people using water supplies closer to home, or switching to at-home toilets. Direct health benefits, namely the value of avoided mortality and/or morbidity, were included in 27 CBAs. Indirect health benefits, such as the time saved by sick patients and their caregivers, were included in 24 studies. Ten studies included averted coping costs, such as expenditure on bottled water or public toilet fees. Three sanitation studies included the market value of biogas or fertiliser replaced through use of urine and compost from faeces (Chuan et al., 2012; Dasgupta et al., 2019; Rodriguez et al., 2011). One study estimated an increase in property value as a one-off benefit at the end of a toilet's useful life (Hutton et al., 2020). Amongst the 26 CBA studies that included both time savings and direct health benefits, the mean contribution of each of them to total benefits was 46% and 41%, respectively (Supplementary Material H).

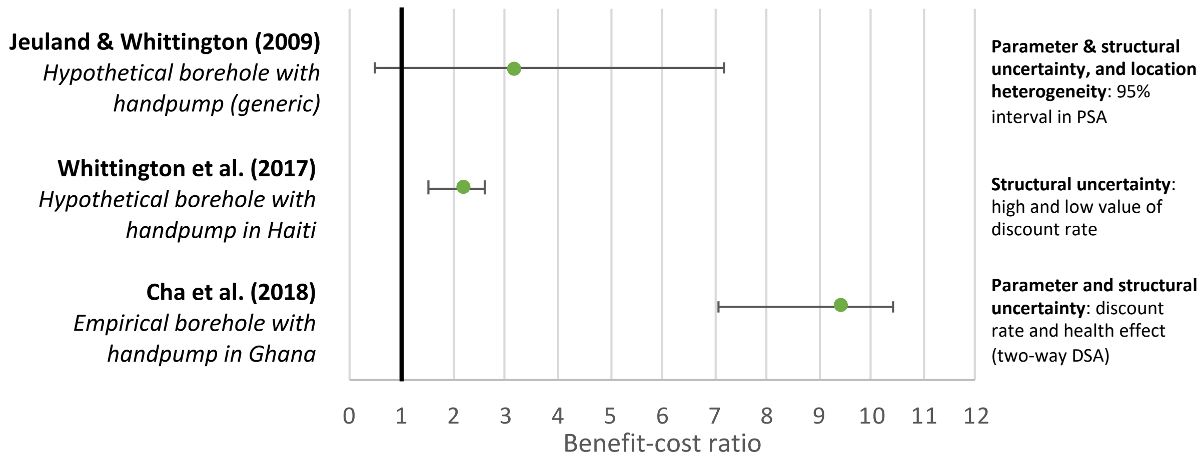
No studies directly valued user-reported quality of life outcomes, such as safety, dignity, convenience and social status for both services, and also privacy and reduced disgust for sanitation (Novotný et al., 2018; Sclar et al., 2018; Stevenson et al., 2012; Wutich and Ragsdale, 2008). Such benefits are not incorporated in CBAs because methods for their measurement and valuation remain absent. Some studies measured these outcomes qualitatively by asking survey respondents about levels of satisfaction (Hutton et al., 2014) or extent of agreement with statements about benefits of owning a toilet (Hutton et al., 2020). For water supply, the value of non-health "aesthetic" benefits has been estimated indirectly, via assumptions about the structure of the demand curve for water, which allows calculation of the residual surplus not captured by health and time savings benefits (Jeuland & Whittington, 2009; Whittington et al., 2017). This approach remains speculative owing to the lack of evidence about the shape of the demand curve, which likely varies substantially across LMIC settings.

Comparison of study results

Only two intervention types were evaluated by at least three fully-costed studies rated “high” or above: (i) communal boreholes with handpumps in rural areas; (ii) rural sanitation campaigns with limited or zero subsidy. In this section, we compare evidence for these two intervention types, and then provide narrative on other common intervention types.

Figure 4: Benefit-cost ratios of communal borehole with handpump interventions in rural areas

Each green marker is the base case BCR for the study, with black error bars denoting its range for the type of uncertainty described to the right. The 95% interval of the BCR in PSA is shown wherever available, with the largest source of uncertainty in DSA as the second choice. Whittington et al. (2009) is not shown, as it presents the same model and result as Jeuland & Whittington (2009). BCR = Benefit-cost ratio; DSA = deterministic sensitivity analysis; PSA = probabilistic sensitivity analysis.

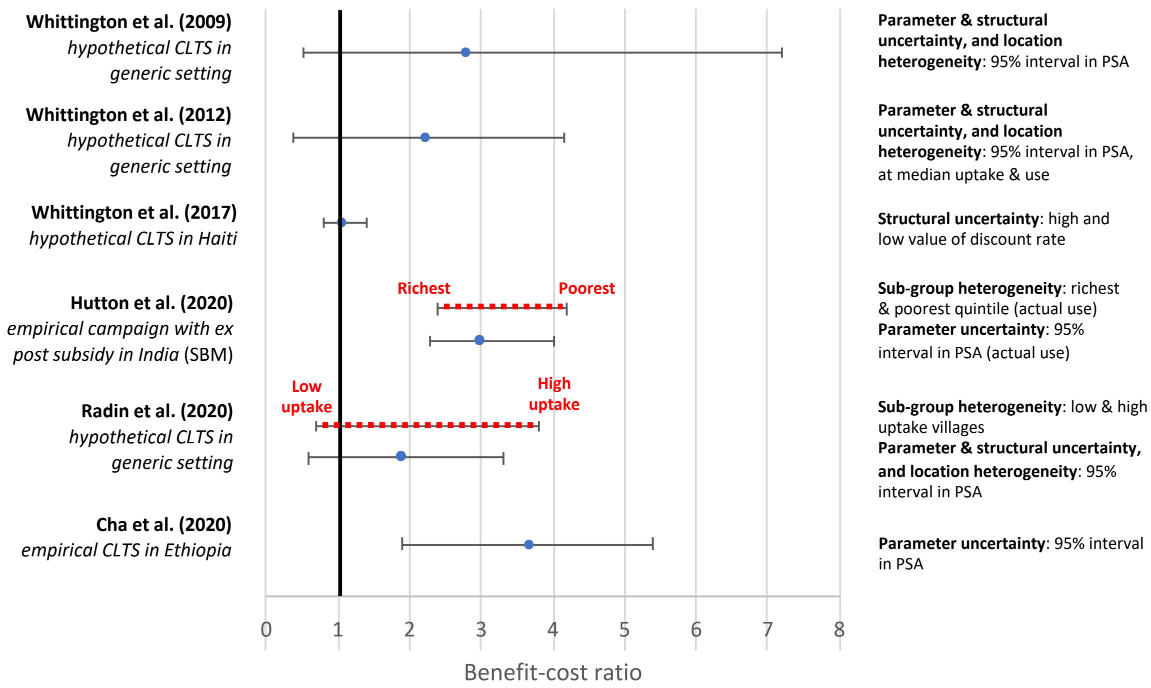


Communal boreholes with handpump in rural areas

Four studies rated “high” or above evaluated fully-costed interventions (Table 1) providing a communal borehole equipped with a handpump. All were CBA studies, and only one evaluated an empirical intervention (Cha et al., 2018). All compared the intervention to use of unimproved water sources (Tables 4-7 in Supplementary Material A). Base case BCRs range from 2.2 – 9.4 (Figure 4). The Cha et al. (2018) study has a high BCR of 9.4, with travel time savings comprising 68% of benefits, and gained work or caregiver days a further 13%. Their result is therefore very sensitive to the value of time, which the authors valued at 100% of the mean wage in the setting. Applying 50% instead, as advocated in guidance (Whittington and Cook, 2018), would bring the base case BCR to around 5 and close to the other studies. From these studies, BCRs between 2 and 5 appear plausible for rural borehole interventions, with low likelihood of BCRs less than 1 if appropriately implemented and targeted.

Figure 5: Benefit-cost ratios of rural sanitation campaigns

Each blue marker is the base case BCR for the study, with black error bars denoting its range for the type of uncertainty described to the right. The 95% interval of the BCR in PSA is shown wherever available, with the largest source of uncertainty in DSA as the second choice. Dotted red bars denote the BCR range for the type of sub-group heterogeneity described to the right. BCR = Benefit-cost ratio; DSA = deterministic sensitivity analysis; PSA = probabilistic sensitivity analysis; SBM = Swachh Bharat Mission; CLTS = community-led total sanitation.



Rural sanitation campaigns

Eight studies rated "high" or above evaluated fully-costed rural sanitation campaign interventions. All were CBAs and two evaluated empirical interventions (Cha et al., 2020; Hutton et al., 2020). All compared the intervention to open defecation (Tables 4-7 in Supplementary Material A), and valued similar types of benefits. Base case BCRs have typically ranged between 1 and 3 (Figure 5), but plausible scenarios across multiple studies yield BCRs below 1. Two studies assessed heterogeneity in efficiency between population sub-groups, finding for example higher BCRs amongst the poorest quintile than the richest (Hutton et al., 2020). Overall, BCRs between 1 and 3 appear plausible for rural sanitation campaigns, if appropriately implemented and targeted.

Other intervention types

For other intervention types, there were not enough studies that met our normative threshold for comparison. For piped water, there are only two fully-costed studies in a specific setting; moreover, their results are not easily comparable due to differences in outcome metrics and quality. One evaluates spring-fed gravity-piped public taps in rural Ethiopia, reporting no BCR (Eklund and Herrmann, 1991), and the other evaluates piped networks with unclear sources serving public taps and household connections in rural India (Fahimuddin, 2012). No CBA studies examine extending piped network provision in a specific urban setting, though one evaluated an intervention which upgraded bulk supply of an existing network to reduce water supply intermittency (Burt et al., 2018).

For household water treatment (HWT), CEAs are more common, with three *ex post* studies of real interventions (Dupas et al., 2020; Rogers et al., 2019; Shrestha et al., 2006). In two studies comparing multiple HWT options, chlorination has lower (more favourable) cost-effectiveness ratios than flocculation with disinfection and ceramic filters (Clasen et al., 2007; Rogers et al., 2019). The only CBAs of HWT were hypothetical studies of biosand filters (Jeuland and Whittington, 2009) or chlorination alongside water supply (Hutton et al., 2007a).

For sanitation interventions other than campaigns, a six-country study in East Asia (Hutton et al., 2014) reported a variety of fully-costed *ex post* analyses of different programmatic approaches implemented in 47 sites (Chuan et al., 2012; Heng et al., 2012; Nguyen et al., 2012; Rodriguez et al., 2013, 2011; Winara et al., 2011). However, results are largely reported for technologies rather than intervention types, meaning that it is not possible to draw conclusions about which types of interventions to encourage uptake of those technologies might be more or less efficient (Tables 2-3 in Supplementary Material A).

Discussion

This systematic review identified 53 studies of the cost-benefit (n=38) and cost-effectiveness (n=22) of water supply (n=30) household water treatment (n=12) and sanitation (n=34) interventions in LMICs. We found study quality to be "high" or "very high" for 19 studies using the CHEERS criteria. We identified 32 studies not included in an earlier, non-systematic review (Hutton and Chase, 2016).

Of the 53 included studies, 20 reported evaluations of empirical interventions, of which seven used outcome data collected as part of an impact evaluation. While hypothetical analyses can be useful to illustrate possible scenarios, the relative lack of empirical analyses is concerning, given the large number of WASH impact evaluations conducted in the past decade (Cumming et al., 2019). Effectiveness is being investigated but allocative efficiency is not. Attaching an economic evaluation to a planned impact evaluation is very common in the health sector (Pitt et al., 2016). Such approaches can have a low incremental research cost, and provide high-quality estimates if planned from the start with cost data collected prospectively. It would do a lot for the WASH evidence base if funders of impact evaluations made such analyses a requirement. There are good WASH examples to build from, both randomised trials (Cha et al., 2020, 2018) and other study designs (Burt et al., 2018; Meddings et al., 2004). Either CBA or CEA might be used, depending on the decision to be informed. However, since WASH interventions deliver diverse non-health benefits, CBA may be the more appropriate choice in most cases (Briscoe, 1984). This does brings increased challenges for outcome measurement and valuation (Robinson et al., 2019a) and may partly explain why WASH economic evaluations are less routine than in health (Boardman et al., 2018).

Few studies evaluated the incremental costs and outcomes of an intervention over another feasible intervention, as opposed to 'no intervention'. It is understandable that such analyses are less common than in the health sector, where interventions are routinely evaluated as additions to packages constantly on offer (e.g. maternity care) or rolled out on a regular basis (e.g. vaccination). WASH interventions, by contrast, are usually delivered very irregularly to a given population even though the resulting technologies and services are subsequently constantly available. Nonetheless, the choice of comparator in WASH economic evaluations is an important one. Baseline conditions can make a big difference to whether an intervention is efficient. “No intervention” never means “no behaviour/technology” since people are always getting water from somewhere and defecating somewhere. A study in rural Laos found that the BCR of moving from open defecation (OD) to a pour-flush pit latrine was 7.8 (Rodriguez et al., 2013, p. 68), while the BCR of moving to the same level of service but from an existing simple pit latrine was 0.9 (*ibid.*, p.71). Modelled interventions and comparators should reflect the decision problem faced by policy-makers; comparing a new intervention to OD in a setting where OD is uncommon is unhelpful.

Two major uses of economic evaluation evidence are to advocate for more resources, and to compare competing options. WASH economic evaluations to date appear to have focused primarily on the first of these uses. Informing common decision problems facing LMIC WASH professionals should be a priority for future research. Such decisions are normally between multiple possible uses of a budget rather than doing something versus doing nothing. Furthermore, CBAs would ideally focus not only on BCR outcomes, but also on the relative size and timing of net benefits and upfront costs. An intervention with a $200,000 return on a $100,000 investment has the same BCR as a $2 million return on a $1 million investment, but the net benefits and likely upfront costs are much larger under the second option. This also has implications for affordability of interventions to governments (Bilinski et al., 2017). The study by Dwumfour-Asare et al. (2020) provides a good example of such a discussion.

Most studies fell short of “high” reporting quality. The increase in numbers of economic evaluations published has not been matched by an increase in their quality (Figure 2). We perceive some quality issues as particularly important. First, simplistic approaches to cost estimation were common in studies we reviewed, despite the fact that costs can contribute just as much to uncertainty as outcomes (Evans and Popova, 2016). Many studies used excessively basic assumptions (such as a percentage mark-up for capital “software”), did not specify data sources, or used aggregated secondary data without discussing which costs might be included or excluded. Recurrent costs can be substantial over the useful life of an asset, but methods for estimating them were often given far less attention than for capital costs. Second, studies would benefit from better characterisation of uncertainty. Only 15% of studies used probabilistic approaches, long the norm in the health sector, which frame decision problems more realistically than deterministic sensitivity analysis. Third, studies paid too little attention to parameters which are the source of substantial uncertainty. Programmes vary in their achievement of uptake and sustained use of services (Garn et al., 2017), but too few studies incorporated these parameters into models. Despite the pre-eminent focus on health effects in many impact evaluations of WASH interventions, this parameter is often far from having the most influence on CBA results (Jeuland and Whittington, 2009). Fourth, transparent presentation of model structure, referencing of input parameters, and tabulation of mean values of costs and benefits disaggregated by type, should all be straightforward. However, too few studies presented them, hiding potentially inappropriate assumptions from scrutiny. Poor transparency also hampers reproducibility (McManus et al., 2019), which is recommended in reference case guidance for both CBA and CEA (Robinson et al., 2019a; Wilkinson et al., 2016). With the general availability of online supplementary materials and repositories for data and methods, there is no longer an excuse for lack of transparency.

Our comparison of studies of similar interventions suggested that BCRs between 2 and 5 appeared realistic for rural borehole interventions, and BCRs between 1 and 3 appeared realistic for rural sanitation campaigns. However, uncertainty and heterogeneity of various kinds mean that BCRs below 1 remain plausible for both these intervention types if they are not targeted to areas where sufficient benefits are likely. These two intervention types – only really applicable to rural areas – comprise only a narrow part of the broad suite of interventions available to WASH decision-makers. Insufficient evidence was available to draw conclusions about other important intervention types. For water supply, the absence of empirical CBAs of piped water connections was surprising. Comparisons of interventions providing off-plot public taps versus on-plot private connections under different scenarios would be useful for informing urban water service provider decisions. For sanitation, more urban studies would be beneficial, particularly those incorporating the full range of costs and benefits arising from waste containment to emptying, treatment and disposal.

Travel time savings were the benefit most commonly measured and valued, followed by direct and indirect health benefits. Next were any averted coping costs not included in these categories (e.g. bottled water purchase during intermittency). Time savings comprised on average a higher proportion of total benefits than direct health benefits, amongst studies that measured both. More empirical stated or revealed preference studies of the value of saved time would therefore be valuable, particularly for sanitation, to reduce structural uncertainty around the value of time saved (Whittington and Cook, 2018). No studies valued quality of life outcomes (e.g. safety, dignity, privacy) either in welfarist or extra-welfarist terms. These benefits are frequently reported as important by users (Gross and Günther, 2014; Jenkins and Curtis, 2005; Jenkins and Scott, 2007; Mukherjee, 2001). More attention has been given to measuring such outcomes in recent years. Examples of new measures include the women's sanitation insecurity profile (Caruso et al., 2017a) and household water insecurity scale (Young et al., 2019). However, these measures employ equal weighting of attributes, and the sanitation insecurity measure lacks an overall score. These features preclude their use in economic evaluation, which would require user-derived valuation (weighting) to incorporate trade-offs between attributes (Brazier et al., 2016). For quality of life benefits to be included in CBA, psychometric measures which meet these criteria are needed, as well as willingness-to-pay studies to enable monetary valuation. Since these quality of life benefits are important to users, excluding them may lead to inefficient or inequitable resource allocation.

To our knowledge, ours is the first systematic review of economic evaluations of WASH interventions. One limitation of our review is that, as others have done, we used the CHEERS reporting checklist as a proxy for study quality and weighted items equally. Some studies may have been conducted to a high standard, but scored poorly in our analysis because they did not report their methods or results with sufficient transparency. Likewise, a small number of studies which scored highly – because their reporting was very transparent – were conducted in ways that we do not judge rigorous. Nonetheless CHEERS provided a useful framework and, overall, scores strongly aligned with our subjective opinion of quality. Second, the focus on full economic evaluations limited the breadth of our review. The per capita cost of an intervention is a cost metric, which is useful for full economic evaluations and for planning purposes (Vassall et al., 2017). However, costs alone cannot inform allocative efficiency, that is, advise decision-makers how to maximise the societal value generated by their budgets. Likewise, valuation studies represent partial economic evaluations on the benefits side, but they cannot directly inform allocative efficiency. Future reviews in these areas would be valuable. Finally, despite systematic searches we may have omitted some studies which met our inclusion criteria. We hope that this review will be updated periodically and welcome recommendations on studies meeting our inclusion criteria (noting those excluded, Supplementary Material D).

Conclusion

We identified evidence that drinking water and sanitation interventions in general can deliver net benefits in many settings; however, benefit-cost ratios are typically modest, indicating that economic evaluation remains important for guiding resource allocation decisions in the WASH sector. Heterogeneity was identified as an important consideration; an intervention which appears efficient at the mean may be inefficient in certain populations and settings. Only a small proportion of existing studies scored “very high” based on the CHEERS criteria, and we recommend that future studies follow recent reference case guidance (Robinson et al., 2019a; Wilkinson et al., 2016) as well as CHEERS reporting guidance. Across all WASH intervention types, more studies are required of empirical interventions which were actually implemented. Future studies could be better tailored to decision problems regularly challenging WASH professionals in LMICs, for example by assessing multiple alternative interventions, and ensuring that comparators represent existing practice.

Funding. This work was supported by the Economic and Social Research Council through a PhD studentship, and the UK Department for International Development. The funders had no role in the identification, design, conduct, or reporting of the analysis.

Conflict of interest. None declared.

Acknowledgements. We are grateful to the experts who proposed studies for inclusion.

Authors' contributions. IR developed the protocol, with inputs from OC, LW, GG, BE and MJ. IR and LW piloted data extraction, and IR undertook the searches. IR and LW undertook screening for eligibility. IR undertook data extraction and quality scoring, with LW independently verifying these for 20% of studies. IR undertook data analysis and wrote the manuscript. All authors contributed to refining drafts

Chapter 5: Research Paper 2 - Qualitative study

Chapter introduction

In Chapter 4 I demonstrated through a systematic review that the user-reported benefits of sanitation related to quality of life, such as privacy, safety and dignity, have never yet been included in an economic evaluation of a sanitation intervention. This is due to the absence of methods for measurement and valuation of these benefits. In Chapter 2 I identified a psychometric measure of sanitation insecurity amongst women as being the only attempt to date to measure aspects of the user-reported experience of sanitation (Caruso et al., 2017a). As a profile measure with seven factor scales and no summative score, the women's sanitation insecurity profile cannot be used in economic evaluation (Brazier et al., 2016).

There is therefore a gap in knowledge on measures to capture quality of life benefits of sanitation which can be used in economic evaluation. In the study reported in this chapter (Research Paper 2), I present results from my qualitative study which was the first step in addressing these gaps. The aim of this qualitative study is to investigate what people most value about sanitation in low-income areas of Maputo, Mozambique, to inform a definition and conceptual model of sanitation-related quality of life. It contributes to objective 2 of this thesis: to develop a psychometric measure of quality of life related to sanitation, based on qualitative research. While informed by this overarching objective, the qualitative study was a standalone piece of research, and the measure development process was separate and subsequent, as reported in Chapter 6. The manuscript was published open access by *Social Science and Medicine* (published by Elsevier) in January 2021, and it follows the standard structure for biomedical journals: introduction, methods, results and discussion. The accepted version is reproduced here. Supplementary Materials are in Appendix B at the end of this thesis.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1701671	**Title**	Mr
First Name(s)	Ian		
Surname/Family Name	Ross		
Thesis Title	Measuring and valuing quality of life in the economic evaluation of sanitation interventions		
Primary Supervisor	Giulia Greco		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Social Science and Medicine (as gold open access)		
When was the work published?	January 2021		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	No	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceptualised the study, developed the methods, and coordinated inputs from the supervisory team. I trained the fieldwork team and participated in piloting and data collection. I conducted the analysis and write-up, revising drafts on the basis of co-author feedback.

SECTION E

Student Signature	[redacted]
Date	19/02/21

Supervisor Signature	[redacted]
Date	19/02/21

How does sanitation influence people's quality of life? Qualitative research in low-income areas of Maputo, Mozambique

Authors:

Ian Ross[1], Oliver Cumming[1], Robert Dreibelbis[1], Zaida Adriano[2], Rassul Nala[3], Giulia Greco[1]

Affiliations:

[1]London School of Hygiene and Tropical Medicine, UK;

[2]WE Consult, Mozambique;

[3]Instituto Nacional de Saúde, Mozambique;

Corresponding author:

Ian Ross

LSHTM, 15-17 Tavistock Place, London, WC1H 9SH

Email: ian.ross@lshtm.ac.uk

Phone: 020 7636 8636

Abstract

Preventing infectious disease has often been the primary rationale for public investment in sanitation. However, broader aspects of sanitation such as privacy and safety are important to users across settings, and have been linked to mental wellbeing. The aim of this study is to investigate what people most value about sanitation in low-income areas of Maputo, Mozambique, to inform a definition and conceptual model of sanitation-related quality of life. Our approach to qualitative research was rooted in economics and applied the capability approach, bringing a focus on what people had reason to value. We undertook 19 in-depth interviews and 8 focus group discussions. After eliciting attributes of "a good life" in general, we used them to structure discussion of what was valuable about sanitation. We applied framework analysis to identify core attributes of sanitation-related quality of life, and used pile-sorting and triad exercises to triangulate findings on attributes' relative importance. The five core attributes identified were health, disgust, shame, safety, and privacy. We present a conceptual model illustrating how sanitation interventions might improve quality of life via changes in these attributes, and how changes are likely to be moderated by conversion factors (e.g. individual and environmental characteristics). The five capability-based attributes are consistent with those identified in studies of sanitation-related insecurity, stress and motives in both rural and urban areas, which is supportive of theoretical generalisability. Since two people might experience the same toilet or level of sanitation service differently, quality of life effects of interventions may be heterogeneous. Future evaluations of sanitation interventions should consider how changes in quality of life might be captured.

Key words: health economics, sanitation, quality of life, qualitative research, toilets, capability approach

Introduction

Preventing infectious disease has often been the primary rationale for public investment in sanitation, defined as the separation of human excreta from human contact (WHO, 2018). Approximately two billion people globally lack access to "basic" sanitation services, defined as an improved type of facility which is not shared with other households (UNICEF / WHO, 2019). An estimated 432,000 annual deaths from diarrhoeal disease are attributable to inadequate sanitation (Prüss-Ustün et al., 2019).

However, health is more than the absence of disease. It is "a state of complete physical, mental, and social well-being" (WHO, 1948). Sanitation also affects these broader aspects of health. A systematic review of the relationship between sanitation and mental well-being identified privacy and safety as root dimensions, predominantly based on qualitative studies (Sclar et al., 2018). Aspects underlying these dimensions were identified as shame, anxiety, fear, assault, dignity and embarrassment. Beyond mental well-being, cleanliness and convenience are also commonly reported as important by users (Novotný et al., 2018). Collectively, we denote these aspects emphasised by users as "quality of life attributes". They are rarely measured in impact evaluations of sanitation programmes, which predominantly focus on disease (Wolf et al., 2018) and toilet use (Garn et al., 2017).

Sanitation cost-benefit studies have noted that improvement in quality of life (QoL) attributes would comprise an economic benefit with a monetary value, but that methods for incorporating them are lacking (Hutton et al., 2020, 2014; Whittington et al., 2020). Economists, then, see privacy, safety or dignity as *outcomes* with an economic value to individuals, but have not attempted to measure them. Sanitation-focused research from other disciplinary perspectives has approached these issues in other ways. The most widely-cited work on QoL attributes has studied them as motives, for example by aiming to identify the behavioural drivers of open defecation (Jenkins and Curtis, 2005; Jenkins and Scott, 2007; Mukherjee, 2001). With disciplinary roots in psychology, these studies see safety for example as a driver of a decision, rather than as an outcome which different interventions might improve to different degrees (Aunger and Curtis, 2013).

Another group of studies assesses QoL attributes as sources of stress and insecurity, often with an epidemiological orientation focused on mental wellbeing outcomes (Caruso et al., 2017b; Kwiringira et al., 2014; Sahoo et al., 2015; Shiras et al., 2018). For example, a measure of women's sanitation insecurity includes aspects of privacy, safety, and so on (Caruso et al., 2017a). It has been used in evaluative studies as a risk factor (Caruso et al., 2018) or effect moderator (Delea et al., 2019) on the causal pathway to mental well-being. In other words, women's sanitation insecurity is applied as an explanatory variable, rather than an outcome in itself.

In the sanitation sector, then, QoL attributes have not been seen as outcomes to be measured and valued, but this is not the case in other sectors. Health-related QoL, for example, is routinely measured in health impact evaluations and applied in cost-effectiveness studies (Karimi and Brazier, 2016). Economic approaches to conceptualising QoL often have the ultimate aim of informing the allocation of public funds, typically leading them to be broadly framed so as to apply to the general population (Fayers and Machin, 2015). Applying an economic perspective to the impact of sanitation on QoL requires a focus on value, or what is important to people. There are divergent traditions within economics on the conceptualisation of value. In utilitarian welfare economics, value is defined by an individual's subjective utility, or the satisfaction they derive from goods or activities (Stiglitz et al., 2009). The capability approach to welfare economics, meanwhile, considers utility problematic due to its focus on individuals' psychological states which can adapt to experience and expectations (Sen, 1980, 1993). Under the capability approach, a good life comprises what people are *able* to do and to be, with QoL attributes identified by discussion of what people "have reason to value" (Sen, 1999). "Conversion factors" are the degree to which an individual can convert "commodities" into capabilities (Robeyns, 2005). An individual has the choice of whether to actually act on those capabilities as "functionings", making the evaluative space an individual's capability to function.

Therefore, while a utilitarian approach to sanitation-related QoL might focus on satisfaction, a capability approach would focus on what people are able to be and do with respect to their sanitation practices. This frames sanitation facilities and services as commodities in support of sanitation-related capabilities. We know of only one peer-reviewed paper applying capabilities to sanitation, a review rather than an empirical study (Barrington et al., 2017). Health economic studies are increasingly using the capability approach to inform the development of outcome measures, based on qualitative research (Al-Janabi et al., 2012; Canaway et al., 2017; Greco et al., 2015; Grewal et al., 2006; Kinghorn et al., 2014).

In assessing the impact of sanitation on QoL, all sources of value can be considered, not only avoiding negative outcomes such as stress or insecurity. Furthermore, in economic applications the *relative* value of different attributes is important . We know of eight studies which provided or enabled a user-reported ranking of motives, stressors or benefits related to sanitation (Gross and Günther, 2014; Hulland et al., 2015; Jenkins, 1999; Jenkins and Curtis, 2005; Jenkins and Scott, 2007; Kiyu and Hardin, 1993; Lagerkvist et al., 2014; Mukherjee, 2001). Only one was in a predominantly urban setting, which was quantitative and focused on motives for use of a "peepoo" bag, not improved sanitation (Lagerkvist et al., 2014).

The aim of this study is to investigate what people most value about sanitation in a low-income urban setting, to inform a definition and conceptual model of sanitation-related quality of life. We use qualitative methods to examine how sanitation contributes to "a good life", by analysing the accounts of users of different types of shared toilet facilities in informal settlements in Maputo, Mozambique. Our underlying objective is to inform the development of a quantitative psychometric measure of sanitation-related QoL.

Methods

We applied a variety of methods to identify and explore attributes of sanitation-related quality of life in the broad evaluative space of capabilities. We used interviews to obtain in-depth accounts, and focus groups to ensure a broader range of views were considered, as well as to engender the deliberation encouraged by the capability approach. We also used pile-sorting and triads, which are structured data collection approaches from cognitive anthropology, to triangulate findings on attributes' relative importance (Weller and

Romney, 1988). We followed the consolidated criteria for reporting qualitative research (COREQ) guidance (Tong et al., 2007), with reporting summarised in Supplementary Material A.

Study setting and intervention

This study was linked to the Maputo Sanitation trial (MapSan), which evaluated an urban shared sanitation intervention (clinicaltrials.gov registration: NCT02362932) (Knee et al., 2020). We used the trial as a vehicle, but our study is not an evaluation of the intervention. In Mozambique, 48% of urban residents lack access to basic sanitation (UNICEF / WHO, 2019). Maputo City, Mozambique's capital, has a population of 1.1 million people (INE, 2019) of which 70% live in informal settlements (UN-HABITAT, 2010). Non-sewered sanitation facilities are used by 89% of Maputo residents, and only 26% of fecal waste in the city is safely managed (Hawkins and Muximpua, 2015). The MapSan trial took place in the low-income *bairros* (neighbourhoods) of the Nhlamankulu district, where multi-household "compounds" with a single entrance to a shared courtyard are common.

People carry out their sanitation practices in a part of the compound called the *casa de banho.* We translate this as "toilet", in the international english sense of meaning any sanitation facility. When referring to specific technologies, we denote those *with* a water-seal ("U-bend") as pour-flush toilets and those *without* as pit latrines (Cairncross and Feachem, 1983). We considered the scope of sanitation practices as perceived by participants, noting that the *casa de banho* (toilet) space is used for bathing and menstrual hygiene management in addition to defecation and urination (Shiras et al., 2018). Before the intervention evaluated by MapSan, toilets mostly comprised an informally-fenced space containing a pit latrine, shared with other households on the compound. Most pit latrines were "traditional", with a soil floor (photos in Supplementary Material B). The intervention was delivered during 2015-16 by Water and Sanitation for the Urban Poor (WSUP), an international non-governmental organisation (NGO), with users making a 10-15% financial contribution. MapSan intervention compounds were provided with a flush or pour-flush toilet discharging to a septic tank, shared with other households on the compound as before. There were two superstructure designs depending on the number of users, all stand-alone buildings not connected to any house. Control compounds continued to use shared pit latrines. We provide more information in about the intervention in Supplementary Material B in the TIDieR checklist format (Hoffmann et al., 2014).

Field team and sampling strategy

The field team comprised four interviewers aged 25-40, two male and two female, led by ZA. All interviewers spoke fluent Portuguese and Changana, the first and second most commonly spoken languages in Maputo. Interviewers underwent a week of training and piloting with IR and ZA which covered informed consent, interview techniques, and transcription skills. Each interviewer undertook interviews and focus group discussions in the setting, observed by IR and ZA, followed by daily team debriefings. Interviewers were from various parts of Maputo and none were known to participants.

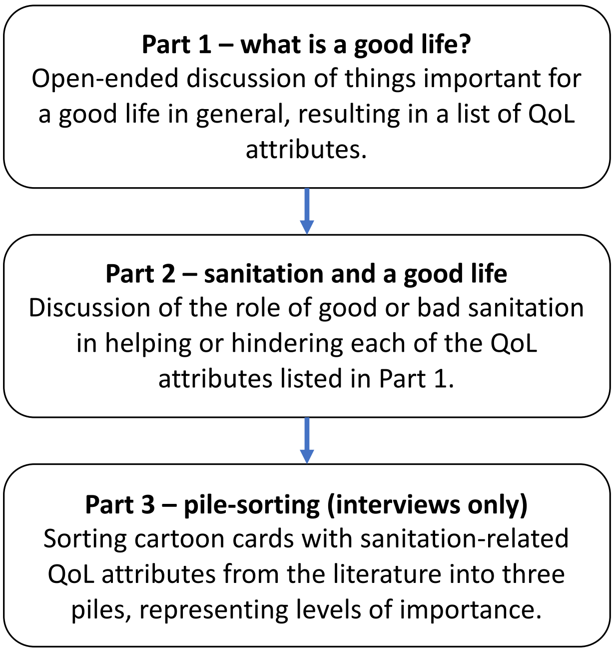
During November-December 2018 we conducted 19 interviews and eight single-sex focus groups of 4-8 participants. The same sampling strategy was followed for interviews and focus groups. To limit respondent fatigue amongst the MapSan study population who had already participated in other trial-related research, we recruited from its broader target population. Specifically, this was multi-household compounds in the same bairros, some of whom had received the same NGO intervention. This meant our study population used the same toilet types as the MapSan intervention and control groups. Our sampling strategy was stratified by three characteristics. We aimed for approximately equal numbers of female and male participants, as well as a mix of respondents by age (18-24, 25-59 and 60+) and toilet type used. We recruited participants by going door-to-door, based on the NGO's records of multi-household compounds, and sampled purposively until a mix of people across strata was achieved. Due to the relatively small sample size, this strategy did not aim to enable exploration of differences between sub-groups, but rather to ensure findings were influenced by a breadth of experiences. Focus groups were convened by the gender and age strata – one was exclusively women aged 18-24, and so on. The majority of interviews and focus groups took place in Portuguese. Changana was used, mostly in short sections, in 16% of transcripts. Interviews and focus groups were audio-recorded, transcribed in Portuguese and translated into English. Interviewers emphasised to participants that they were independent researchers and not working for the implementing NGO or local government.

Data generation and topic guides

One interviewer of the same gender as the participant carried out each interview, which lasted on average 40 minutes. The interview topic guide, adapted iteratively during training and piloting, is included in Supplementary Material C and summarised in Figure 1. Part one aimed to identify valued attributes of a good and bad life in general. Part two discussed the role of sanitation in affecting those valued attributes (Figure 1) considering all sanitation practices important to participants. Part three used pile-sorting (Weller and Romney, 1988), which generated structured data rather than a verbal account. Participants were presented with 15 cards, each with an attribute of good sanitation identified in previous reviews (Jenkins and Sugden, 2006; Novotný et al., 2018). Cards comprise the bar labels in Figure 4 and included cartoon depictions (Supplementary Material C). Participants were asked to choose the five that they thought were most important for a good life, then a second set of five for the next level of importance. The remaining five cards comprised the third set.

Focus groups were managed by one moderator and one notetaker, of the same gender as participants. Each lasted on average 60 minutes. The first six focus groups followed the same guide as interviews, but omitting the pile-sorting exercise in the interests of time and practicality. For the final two focus groups (one male, one female), we reconvened 6-8 participants from previous focus groups and interviews. First, emerging findings were presented and reflections sought in a “participant checking” discussion (Green and Thorogood, 2009), then we proceeded with a triad exercise (Weller and Romney, 1988). Participants were presented with three attributes and asked to choose the most important of them, and the process then repeated with different combinations. Attributes on triad cards were based on emerging analysis, and fewer in number than pile-sorting cards to reduce respondent fatigue (Supplementary Material D).

Figure 1: Structure of IDI and FGD topic guide



Data analysis

Value was the "economic lens" (Coast and Jackson, 2017) in design of data generation activities and subsequent analysis, specifically the capability approach and its focus on "reason to value" (Sen, 1999). The primary output of our analysis is a conceptual model comprising theory about what people value about sanitation, alongside contextual qualitative description. The focus on value and capabilities influenced our design and analysis in several ways. First, it meant we prioritised the exploration of the relative importance of attributes through multiple methods. Second, the topic guide was organised such that the evaluative space was broad, focused on "a good life". Third, core attributes in the eventual conceptual model were framed as capabilities. Fourth, we coded passages which related to conversion factors, i.e. the characteristics of individuals and their physical and social environment.

We used 'framework' analysis to interpret transcripts (Ritchie and Spencer, 1994). Framework is not associated with a specific epistemological position, and allows for both inductive and deductive coding (Green and Thorogood, 2009; Ritchie and Lewis, 2003). We followed an interpretive approach, acknowledging that data based on interaction in interviews and focus groups only describes one version of reality (Coast et al., 2017). We followed the five steps of framework analysis (Ritchie et al., 2003) to understand the data: (i) familiarisation; (ii) identifying a thematic framework; (iii) indexing; (iv) charting; and (v) mapping and interpretation. As part of step 1, after ZA had verified transcription and translation, IR uploaded them into nVivo 12 (QSR International, 2018) and wrote a summary memo of each. Transcripts were presented in English and Portuguese paragraph by paragraph, enabling coding of the English version with easy reference to the original Portuguese where necessary. For step 2, coding took an inductive approach, without applying any *a priori* codes. IR, who understands Portuguese, open-coded four interview transcripts as an initial batch and discussed the resulting codes with GG. Closely-related codes were combined and redundant codes deleted. The codebook was applied to the next batch of four, the process repeated, and a framework gradually emerged. The codebook was rarely updated after the third iteration, suggestive of theoretical saturation. The final codebook ("index") was then applied to remaining interviews and all focus groups (step 3). Based on charting (step 4), IR wrote analytical memos for the most salient concepts, in support of interpretation (step 5).

Through this process, we arrived at a set of "core attributes" which are sanitation-related capabilities, and several "underlying concepts" for each. We would highlight three aspects of the process. First, core attributes incorporated both positive and negative aspects of the same concept (e.g. pride and shame) and linked concepts (e.g. smelling faeces and seeing maggots). Second, in refining codes, we built on Al-Janabi's (2012, p. 169) approach, to reflect "less the specific influences on well-being (e.g. work) and more the concepts that could be influenced by multiple factors (e.g. stress). … that represented what was ultimately important in individuals' lives". Third, we aimed to be reflexive in our analytical processes of identifying and applying a framework, considering alternative ways in which concepts could be labelled and codes aggregated. To the extent possible, codes reflected the *in vivo* phrases used by respondents in Portuguese. However, it is impossible to exclude the role of the researcher in shaping analysis.

Core attributes were identified only through the above analysis of transcripts, with other data analysis used only for triangulation. After core attributes were identified, we added intensity codes to interview excerpts coded with those attributes (Saldaña, 2015). We did not do this for focus groups, since the relative tone attached to different passages is more uncertain when many people are contributing and interacting. Passages were coded as "low intensity" when participants mentioned the topic in passing or in a short impassive phrase, "medium" when discussed in more detail or with mild emotion, or "high" when a lengthy example or emotive language was used. We used nVivo coding queries to map intersections of intensity codes with content codes, and tabulated frequencies. In analysis of pile-sorting data, concepts in the top pile were given three points (two in sensitivity analysis), the second pile one point, and the remaining pile zero points.

Ethical considerations

All participants provided written informed consent to participate. We informed participants of their right to end discussions at any time, and all audio recordings were permanently deleted after verification of transcripts. The study received approval from the Research Ethics Committee of the London School of Hygiene and Tropical Medicine (Ref: 14609) and from the *Comité Nacional de Bioética para a Saúde* (IRB00002657) at the Ministry of Health in Mozambique.

Results

There were 73 participants in the study overall, of which 41 were women and 32 men (Table 1). Most participants (75%) used a toilet shared by 2-5 households. There were more users of flush toilets than pit latrines, but the majority of these had received the NGO intervention in the past 1-2 years so could reflect on pre-intervention experience with pit latrines. Two female interview participants were pregnant, and two male interview participants had a physical disability affecting their mobility.

Below we present findings for each of five capabilities which are core attributes of sanitation-related QoL, representing what participants most valued about sanitation in this setting: health, avoiding disgust, avoiding shame, safety, and privacy. We provide illustrative quotations for each (Portuguese in Supplementary Material E). Afterwards, we present a conceptual model illustrating how the attributes fit together, and findings from the triangulation analyses. In participant checking undertaken in the final two focus groups, no concerns or proposals were raised about findings emerging at that time.

Table 1: Demographic characteristics of participants

	IDIs % and (n)	FGDs % and (n)	Overall % and (n)
Gender			
Female	53% (10)	57% (31)	56% (41)
Male	47% (9)	43% (23)	44% (32)
Age			
18-24	37% (7)	37% (20)	37% (27)
25-59	42% (8)	50% (27)	48% (35)
60+	21% (4)	13% (7)	15% (11)
Education			
Did not complete primary	26% (5)	41% (22)	37% (27)
Primary or incomplete secondary	42% (8)	48% (26)	47% (34)
Completed secondary or above	32% (6)	11% (6)	16% (12)
Tenancy			
Owners	79% (15)	76% (41)	77% (56)
Renters	21% (4)	24% (13)	23% (17)
Toilet type*			
NGO-supported flush / pour-flush	47% (9)	17% (9)	25% (18)
Other flush / pour-flush	16% (3)	46% (25)	38% (28)
Pit latrine	37% (7)	37% (20)	37% (27)
Households using toilet			
1	0% (0)	13% (7)	10% (7)
2-5	79% (15)	74% (40)	75% (55)
6+	21% (4)	13% (7)	15% (11)

*As set out in the methods section, we recruited from the target population of the MapSan trial, not its study population. Users of "NGO-supported flush / pour-flush" are analogous to the MapSan intervention group, and users of "pit latrine" are analogous to the control group.

Health

The two main concepts underlying the health (*saúde*) attribute were disease (*doença*) and peace of mind (*paz de espírito*), Almost all interview participants (18/19) mentioned one or both, with roughly twice as many mentioning disease. Unspecified disease was most commonly cited, but specific symptoms (diarrhoea, vomiting) or pathogens (cholera) were too. There was general understanding that children touching faeces would spread disease, also emphasising concern for others as a relational aspect of QoL:

> *"It is difficult for us to control what children do. An adult knows they shouldn't touch something, or they'll catch germs, but a child doesn't know." Male interview, 36 (EAGJ04)*

Healthcare expenses as a result of disease were cited as a problem deriving from poor sanitation, and flooding was cited as a risk factor for sanitation-related diseases:

> *"When it rains the faeces in the pit rise up, then we get diseases because cholera comes from there" Female interview, 71 (EANC04)*

The mental wellbeing aspect of health was most commonly framed as "peace of mind", but also as feeling "at ease" (*a vontade*) and "relaxed" (*tranquilo*). It was far more often cited in relation to other attributes (disgust, shame, and privacy in that order) than on its own:

> *"You feel under pressure when the bathroom is dirty, and you don't feel at ease." Female focus group, 18-24 (FGF01)*

> *"Your neighbours will know the origin of the smell and will start to talk about it, and you can't feel relaxed." Female interview, 76 (EAET04)*

Peace of mind was also referred to positively, for example people with better-quality toilets reporting feeling relaxed while using it. It was sometimes mentioned without being linked to another specific concept:

"Having a good toilet contributes positively to all these aspects, mental health, wellbeing for the soul, and general health as a whole." Male interview, 64 (EAGJ03)

Avoiding disgust

Two concepts underlying the disgust (*nojo*) attribute were sights / smells, and feeling clean / touch. Almost all interview participants (17/19) mentioned one or both, with sights/smells cited approximately twice as often. Seeing faeces was a common trigger of disgust for users of pit latrines, as they can often be seen through the drop-hole, but also for users of flush toilets which were not clean. Maggots, cockroaches and flies were also visual triggers of disgust for pit latrine users.

"It is something so horrible to see other people's faeces" Male interview, 19 (EAGJ02)

The toilet's smell was important both at the time of using it and at other times:

"You cannot eat because you lose your appetite … due to the smell. You don't even feel free to come out of your house because it smells bad out there." Female focus group, 25-59 (FGF02)

When considering good toilets, people talked about positive consequences of a lack of disgust, for example being able to do more things in there rather than wanting to rush out immediately:

"[In this toilet] I feel like I am in the kitchen. With no bad smell, it seems like you could even drink tea in there, without realising you are in a toilet." Female interview, 71 (EANC04)

Users of better-quality toilets reported appreciating their cleanliness, illustrating the positive side of this attribute. People would prefer to use a toilet that was clean, but also to *feel* clean both while there and after leaving (for example, after bathing).

“When the house is clean but the toilet is not, this is undignified.” Female focus group, 60+ (FGF03)

Avoiding shame

The two concepts underlying the shame (*vergonha*) attribute were “what others think or say” (pride/status), and dignity. Three quarters of interview participants (15/19) mentioned one, or more, and were roughly twice as likely to mention “what others think or say” as the other.

“A person's toilet becomes the mirror of that person.” Male interview, 64 (EAGJ03)

The most commonly reported trigger for shame was the disdain of others on account of the smell or appearance of the toilet. Shame could also be caused just by knowing that other people could smell the toilet, without even interacting with them:

“Everyone will refer to you according to the state of your toilet, saying ‘it's there at her house that the toilet smells’ … nobody respects you.” Female interview, 27 (EAET05)

“People who go down my road smell the stench from my toilet. Then when they later pass me on the street they will look at me in a different way.” Male interview, 19 (EAGJ02)

Several respondents with ‘good’ toilets reported feeling proud or feeling more respected by visitors, reflecting the positive side of this attribute:

“When I get visitors, I can let the person use the toilet without fear. I think this makes people look at me differently, with respect.” Male interview, 28 (EAJP05)

“If a visitor asks to go to the toilet and sees it in good condition, they’ll say ‘wow, that lady’s house is hygienic’” Female focus group, 60+ (FGF03)

Safety

The two concepts underlying the safety (*segurança*) attribute were accidents / falls, and violence (physical or sexual) which included “peeping”. Two thirds of interview participants (13/19) mentioned one or both, with accidents / falls reported roughly twice as frequently. For safety in general, prospective concern about things which *might* happen appeared much more prevalent than actual experience.

Respondents identified these risks with respect to themselves but also family members, again emphasising the relational aspect of QoL. For example, risk of injury was reported in relation to a child falling into the pit, or a poorly-constructed toilet collapsing:

“That toilet built from car tyres is a hazard – when it rains it could come crashing down at any moment.” Male focus group, 25-59 (FGM02)

Respondents with reduced mobility (e.g. pregnant women, older people, disabled people) were more likely to report fearing falling into the pit or falling over while squatting. With low-quality toilets, a risk was not being able to see properly at night:

“I’m afraid to use it at night because I wouldn’t know which way to enter, where to tread inside, and I would be afraid of falling into the hole.” Female interview, 71 (EANC04)

Regarding violence, participants more often described the generalised risk of bandits and thieves. Sexual assault was seen as a risk for both men and women. People with high-quality toilets did not necessarily feel safer at night, because everyone needed to leave their house into an insecure compound to access the toilet building:

> *"There are people who are raped while they use these toilets. … there are times we even have to defecate in a bucket because we fear bandits." Female interview, 27 (EAET05)*

Privacy

The main concepts underlying the privacy (*privacidade*) attribute were being seen and being disturbed. Two thirds of interview participants (12/19) mentioned one or both, with "being seen" roughly four times as likely as being disturbed. Sometimes respondents knew people could see them over or through the walls or door, as they themselves could see passers-by or hear children laughing. However, there was also the fear that an unseen person might be "peeping" (see "safety" above). Privacy was important for all types of sanitation practices, including bathing and menstrual hygiene:

> *"You cannot imagine the gymnastics I do when I have my period. I do not feel relaxed because I do not know if I'm being watched." Female interview, 27 (EAET05)*

From the viewer's perspective, privacy could also be infringed unintentionally:

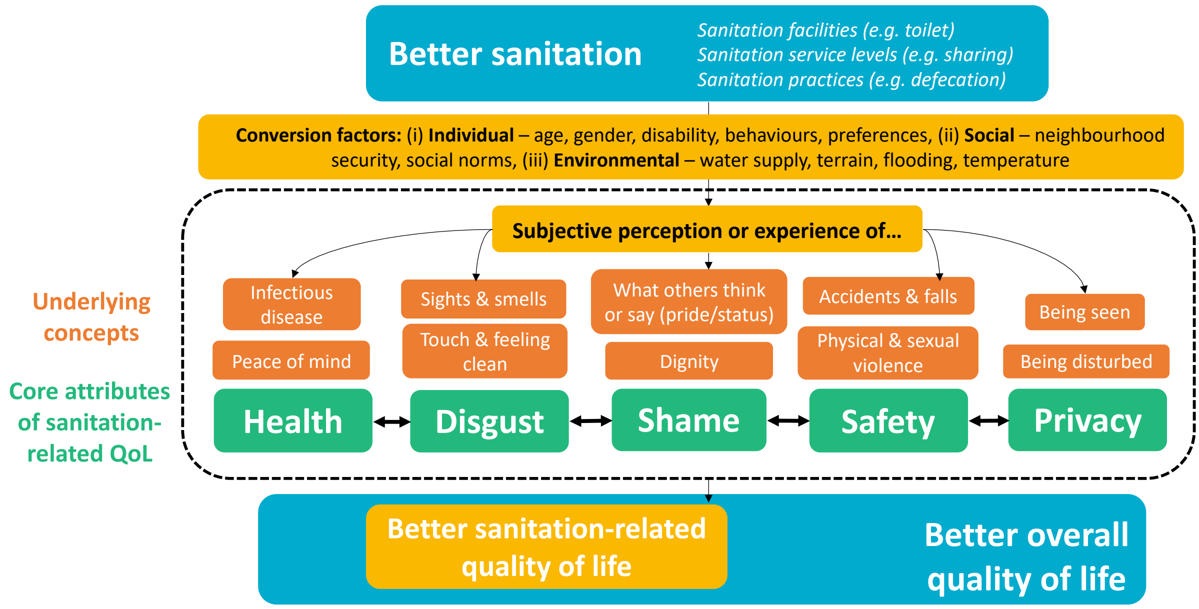
> *"While you walk to work, … you might see a woman with just a bit of capulana [fabric], when she is naked taking a bath." Male interview, 28 (EAJP05)*

The concept of being disturbed concerns someone else entering an unlocked or door-less toilet, without knowing you were inside.

> *"When a bathroom is not secure you do not feel free to use it because at any moment an individual can enter." Female interview, 76 (EAET04)*

Some respondents also mentioned a good-quality private toilet providing a place to do make-up, trim body hair, or be intimate with one's partner through showering together or sex.

Figure 2: Conceptual model for sanitation-related QoL



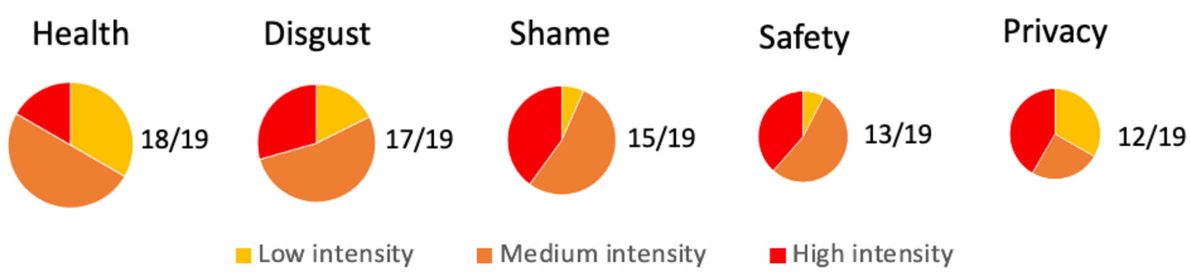
Conceptual model

Figure 2 presents a conceptual model for sanitation-related QoL, which visualises five findings. First, we define sanitation-related QoL as "the subset of overall QoL which is directly affected by sanitation practices or services". This definition draws on analogous definitions of health-related QoL (Karimi and Brazier, 2016; Peasgood et al., 2014). The scope of sanitation practices is as perceived by users, but is assumed to include defecation, urination, menstrual hygiene, and any related practices users consider important. Second, five capabilities were identified as core attributes of sanitation-related quality of life (green in Figure 2). Underlying each is a number of concepts (in orange). Third, an improvement in sanitation facilities, services or practices might cause an improvement in overall QoL, demonstrating the instrumental value of sanitation. That improvement might act via changed experience or perception around one or more of the capability-based attributes. Fourth, any effect of improving sanitation on QoL may be moderated by conversion factors. Examples include respondents with reduced mobility being more likely to fear falling into the pit, or people with good-quality toilets fearing using them at night. Based on our findings, we hypothesise that conversion factors include: (i) individual conversion factors, such as gender, age and disability, (ii) social conversion factors, such as neighbourhood security and social norms, (iii) environmental conversion factors, such as flooding and the level of the water table. Fifth, while these five attributes are distinct sources of value, they are also inter-related. Sometimes respondents' safety concerns arose from privacy deficits, but many also related to the journey to the toilet through an insecure compound. A toilet being dirty or disgusting was perceived as bad regardless of the possible health consequences, which were mostly not mentioned as part of the same point. Shame was related to disgust and privacy, e.g. being embarrassed at using a smelly toilet or being seen.

Relative importance of attributes

Participants discussed attributes with different frequency. In Figure 3, the relative size of pies represents the proportion of interview participants mentioning each attribute at least once. Participants' intensity of speech also varied – pie charts show the highest level of intensity used by each participant mentioning that attribute. This analysis shows that while health and disgust were mentioned most often, a lower proportion of participants mentioned them using highly intense phrases. This contrasts with shame and safety in particular, which were more likely to be mentioned in medium or high intensity language, showing the limitations of analysing frequency alone.

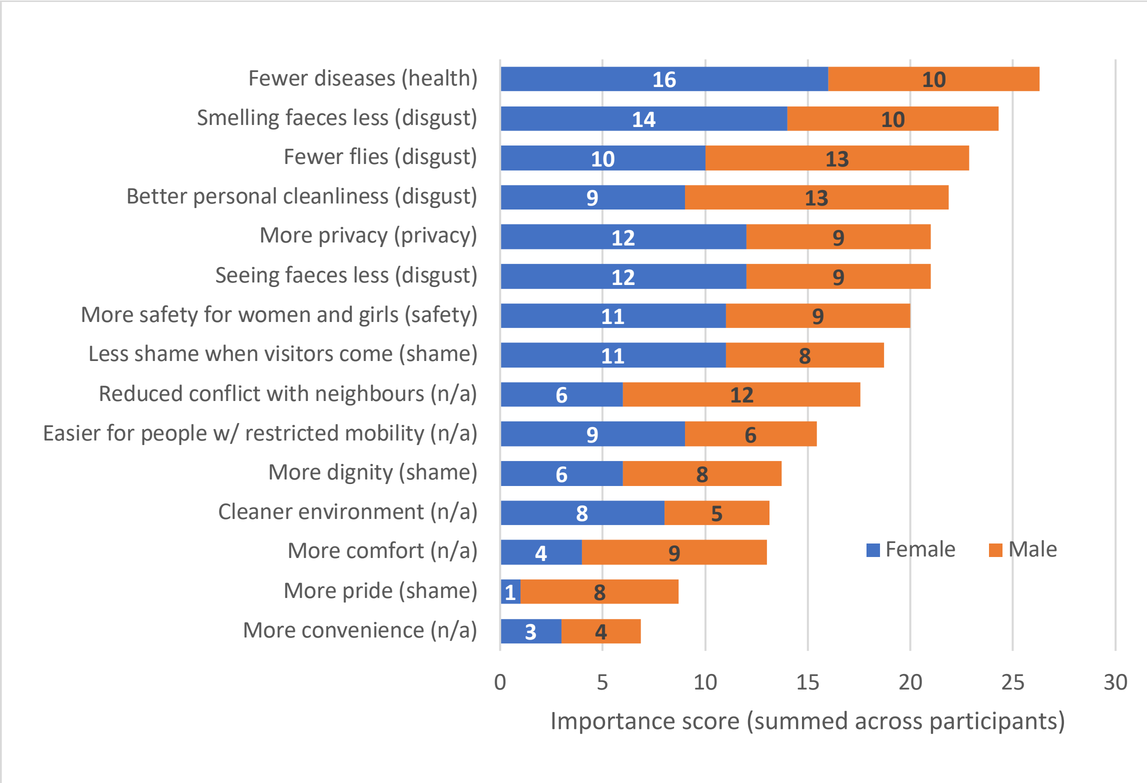
Figure 3: Frequency and intensity with which attributes were mentioned in IDIs



In the pile-sorting findings (Figure 4), each bar shows the total of importance scores for 15 concepts, normalised to account for gender imbalance. Bar labels note the most closely-related core attribute in parentheses. Though the primacy of health and disgust is in line with frequency data presented above, pile-sorting data imply a slightly different ranking for the other attributes: (i) health, (ii) disgust, (iii) privacy, (iv) safety, (v) shame. Sensitivity analysis (two points instead of three for the first pile) did not change these findings.

Figure 4: Pile-sorting attributes during IDIs (n=19)

Labels (without bracketed part) were written on cards in Portuguese and read out by interviewers. n/a = does not map onto any single attribute.



Health and disgust are again first and second in the triad data (Supplementary Material D), but the other attributes are again ordered slightly differently. This indicates uncertainty around the relative value of those three attributes for participants, given the small sample sizes involved. In ascertaining relative value, we would place more weight on the pile-sorting and triad data, where respondents directly traded off attributes. The frequency and intensity data (Figure 3), by contrast, rely on our interpretation. Concepts which do not map onto any single attribute in our model are denoted “n/a” in bar label parentheses (discussed in Supplementary Material D). These concepts scored lower than the highest-ranked concepts linked to the five core attributes, and were in the bottom 40% of cards overall. There were some differences in scoring between sexes. The biggest absolute differences in scores were for pride, reduced conflict with neighbours (both of which men valued more) and fewer diseases (which women valued more). Women were only slightly more likely to value privacy and safety. However, given the small sample size of sub-groups, we would not place much weight on these differences.

Discussion

In this study, we carried out qualitative research into what people most valued about sanitation. We used the findings to develop a definition and conceptual model of sanitation-related quality of life in a low-income urban setting. Using attributes of “a good life” as an entry point for discussion, we identified five capabilities as core attributes of sanitation-related QoL: health, avoiding disgust, privacy, safety and avoiding shame. Our conceptual model outlines how improvements in sanitation commodities might improve capabilities via changes in these five attributes, and how capabilities are moderated by personal and environmental conversion factors.

Three aspects distinguish our study from previous work in this area, particularly studies focused on sanitation-related stress and insecurity (Caruso et al., 2017b; Kwiringira et al., 2014; Sahoo et al., 2015; Shiras et al., 2018). First, we conceptualise sanitation-related QoL as an outcome like health-related QoL, by contrast to insecurity and stress which are usually conceptualised as risk factors affecting mental health outcomes (Caruso et al., 2018). As a result, health is an attribute *within* the concept of sanitation-related QoL (Figure 2). Second, we conceptualise sanitation-related QoL as applicable to the general population, while the literature on insecurity and stress has usually focused exclusively on women. Third, our approach to qualitative research was rooted in economics, in particular value and the capability approach.

We used pile-sorting and triads as alternatives to simple ranking in exploring relative value. The broad concurrence of the top-ranked attributes triangulated across methods gives us confidence in these findings. While concepts on pile-sorting cards were imposed on participants based on the literature, the corresponding advantage was that the participant's choice was direct, rather than the indirect interpretation of transcripts by the researcher.

Applying capability theory about conversion factors helps emphasise how two people achieving an identical improvement in objective toilet quality may experience a dissimilar QoL effect. Two people may experience the same toilet very differently. For example, because of social norms around gender, a middle-aged man might have a higher conversion factor than an adolescent girl, if she has different expectations of privacy and safety to achieve a given level of capability. His conversion factor might also be higher than an older man with restricted mobility who worries more about falling and finds it harder to avoid touching disgusting surfaces. This is important because many sanitation interventions and evaluations focus on access to a given technology or level of service, implicitly assuming it delivers similar benefits to all users.

Environmental conversion factors are also important. The social and environmental context in which a toilet commodity is used affects the capabilities an individual can derive from it. In several previous studies, convenience was identified as a valuable attribute of household toilets, as compared to open defecation or public toilets (Novotný et al., 2018). Our participants did not talk about convenience, since all used on-plot toilets and took this for granted. Similarly, it was unsurprising that water supply was rarely mentioned as an important influence on sanitation capabilities, contrary to some other settings (Sahoo et al., 2015), since all participants had a fairly reliable piped water supply within the compound.

Despite the differences in framing and methods, the five identified attributes are broadly consistent with studies of insecurity, stress and motives related to sanitation in both rural and urban areas, which is supportive of theoretical generalisability (Novotný et al., 2018; Sclar et al., 2018). One area of contrast relates to the ranking of attributes rather than their identification. "Fewer diseases", as the card was framed, was consistently identified as the most valuable in pile-sorting and triads. It was also the most-frequently mentioned concept in interviews. In previous studies, disease prevention (or less often "health", possibly in its broader sense) was typically ranked second or third, behind other attributes (Gross and Günther, 2014; Hulland et al., 2015; Jenkins, 1999; Jenkins and Curtis, 2005; Jenkins and Scott, 2007; Kiyu and Hardin, 1993; Lagerkvist et al., 2014; Mukherjee, 2001). The difference between first place and second or third place is not that great, but debate in the literature on the relative importance of disease prevention (Jenkins and Curtis, 2005) suggests that this merits discussion. We propose three possible explanations for disease prevention being the most valuable attribute of sanitation in this setting. First, all but one of the previous studies was in a predominantly rural setting. Our setting was urban, where populations generally have higher levels of education (Zhang, 2006) and greater exposure to media, which can influence health-related knowledge (Agüero and Bharadwaj, 2014; Yaya et al., 2018). Second, a participant identifying something as valuable or important is different to them identifying what motivates a specific behaviour (Aunger and Curtis, 2013). Therefore, given the majority of these previous studies were motives-oriented, they were measuring something slightly different to value. Third, half our interview sample comprised people who had received an NGO sanitation intervention a year or two previously, which included direct and indirect health messaging. This may have contributed to a real change in the relative value of attributes, or to social desirability bias if interviewers were perceived as linked to the intervention, despite being told this was not the case.

Limitations

Our findings reflect the setting in which the data were generated, namely a low-income area of urban Maputo where shared sanitation was common. This limits the transferability of our findings to other settings to some extent. However, since the attributes identified broadly align with studies in diverse countries under different disciplinary perspectives, a level of theoretical generalisability may be claimed. As with any interpretivist qualitative research based on conversational interaction, our findings describe only one version of reality. The majority of analysis was undertaken by a non-Mozambican researcher (IR) who may have misinterpreted some interaction in transcripts, though important or unclear passages were always discussed with ZA and the field team. The relatively small sample size precluded comparison of findings by sub-groups of gender or age.[4] Likewise, only tentative conclusions can be drawn from the pile-sorting findings, particularly in relation to the differences in relative value attributed by women and men. Reasons for caution in interpreting the frequency data are not only the sample size, but also the fact that topic guides were flexible, with sanitation discussions guided by the QoL attributes respondents had mentioned in part 1. The fact that some participants had received an intervention, and others had not done so but were likely aware of the intervention in their area, may have biased their responses in unpredictable ways. For the 16% of interviews that involved some Changana, some meaning was possibly lost in translation to Portuguese, despite interviewers being fluent in both languages.

Conclusion

Our findings illustrate that people in low-income areas of Maputo, Mozambique, valued many different aspects of sanitation. Our interpretation of their accounts was captured in five core attributes of sanitation-related QoL: health, disgust, privacy, safety and shame. Our intention is to use these findings to inform the development of a quantitative measure in this setting, alongside quantitative methods of attribute valuation. We hope that others might explore sanitation-related QoL in other settings and populations to validate or develop the conceptual model. Sanitation interventions might improve different attributes of sanitation-related QoL to different degrees. The sixth sustainable development goal emphasises that sanitation for all should be adequate and equitable. Since two people might experience the same level of sanitation service very differently, thresholds of adequacy may differ across individuals and QoL effects of intervention may not be equitable. Future evaluations of sanitation interventions should consider how changes in quality of life might be captured, as well as changes in level of service and health outcomes.

[4] This sentence and those following would have benefited from clarification, but this in the final published version, so clarification is provided in this footnote. I intended to explain why there was a single conceptual model across the study population, rather than have different conceptual models for sub-groups. In this chapter I *do* of course compare findings between individuals and sub-groups. An example in the transcript analysis is in noting that pregnant women, older people, and disabled people were more likely to report fearing falling into the pit or falling over while squatting. An example in the pile-sorting results is in presenting and discussing data in Figure 4 by gender.

Funding. 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 interest. None declared.

Acknowledgements. We greatly appreciated the cooperation of interview and focus group participants. We also appreciated the efforts of the inteviewers: Euclimia Titosse, Neusia Chavate, João-Pedro Guambe and Gimo Jenga. Ian Ross benefitted from discussions with Val Curtis and Sarah Bick at an LSHTM writing retreat, participants at the UNC Water and Health Conference 2019, and participants at the Human Development and Capabilities Association conference 2019. Finally, we would like to thank three anonymous peer reviewers whose inputs substantially improved the manuscript.

Author contributions. Ian Ross: Conceptualization, Methodology, Investigation, Formal analysis, Writing - Original Draft. **Oliver Cumming**: Supervision, Writing - Review & Editing, Funding acquisition. **Robert Dreibelbis**: Methodology, Writing - Review & Editing. **Zaida Adriano**: Methodology, Data curation, Investigation, Writing - Review & Editing. **Rassul Nala**: Writing - Review & Editing, Funding acquisition. **Giulia Greco**: Supervision, Methodology, Writing - Review & Editing

Chapter 6: Research Paper 3 - Measure development

Chapter introduction

In Chapter 5, I investigated what people most valued about sanitation in low-income areas of Maputo, Mozambique, and developed a definition and conceptual model of sanitation-related quality of life. That was the first step towards objective 2 of this thesis: to develop a psychometric measure of quality of life related to sanitation, based on qualitative research.

The study reported in this chapter (Research Paper 3) comprises the second part of achieving that objective. The aim of this study is to develop and evaluate a novel measure of sanitation-related quality of life (SanQoL) in urban Maputo, Mozambique. I am proposing to submit the paper to *Quality of Life Research* (published by Springer), but provide more detail in the body of the text here than in the version which would be submitted to the journal. As is relatively common with measure development papers, methods and results for each step in the process are reported concurrently, followed by a separate discussion. Supplementary Materials are in Appendix C at the end of this thesis.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1701671	**Title**	Mr
First Name(s)	Ian		
Surname/Family Name	Ross		
Thesis Title	Measuring and valuing quality of life in the economic evaluation of sanitation interventions		
Primary Supervisor	Giulia Greco		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Quality of Life Research
Please list the paper's authors in the intended authorship order:	Ian Ross, Oliver Cumming, Charles Opondo, Zaida Adriano, Rassul Nala, Joe Brown, Robert Dreibelbis, Giulia Greco
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceptualised the study, developed the methods, and coordinated inputs from the supervisory team. I trained the fieldwork team and participated in piloting and data collection. I conducted the analysis and write-up, revising drafts on the basis of co-author feedback.

SECTION E

Student Signature	[redacted]
Date	24/10/20

Supervisor Signature	[redacted]
Date	24/10/20

Development and validity of a novel measure of sanitation-related quality of life (SanQoL)

Authors:

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

Affiliations:

[1]London School of Hygiene and Tropical Medicine, UK;

[2]University of Oxford, UK;

[3]WE Consult, Mozambique;

[4]Instituto Nacional de Saúde, Mozambique;

[5]Georgia University of Technology, USA;

Corresponding author:

Ian Ross

LSHTM, 15-17 Tavistock Place, London, WC1H 9SH

Email: ian.ross@lshtm.ac.uk

Phone: 020 7636 8636

Abstract

Purpose: To develop a measure of sanitation-related quality of life (SanQoL) in urban Maputo, Mozambique, and assess its validity and reliability.

Methods: After initial qualitative research, development and assessment of the SanQoL index followed four steps: (1) item development, (2) piloting and cognitive interviews, (3) valuation (weighting) of attributes by ranking them on a visual analogue scale, explored using mixed-effects ordered logistic regression, (4) psychometric evaluation, including hypothesis testing using generalised linear mixed models, and assessment of test-retest reliability. A quantitative survey in support of steps 3-4 was undertaken with people living on intervention and control compounds enrolled in the Maputo Sanitation Trial.

Results: Qualitative research identified five attributes within a conceptual model of SanQoL: disgust, health, shame, safety and privacy. In the quantitative dataset (n=424), there was evidence of association for 86% (30/35) of hypothesised presence or absence associations between the five SanQoL attributes and characteristics of users and toilets. Correlation between SanQoL index values and the WHO-5 mental wellbeing index was 0.24, low but positive as hypothesised. Intraclass correlation (ICC) was 0.72 for two respondents using the same toilet. Test-retest reliability was good (ICC: 0.87). There was insufficient evidence for differences in attribute ranks by gender, age or intervention status.

Conclusion: This study provides evidence in support of validity and reliability of the SanQoL measure. Such a measure may help quantify the broader impact of sanitation interventions and better value incremental outcomes in cost-effectiveness and cost-benefit studies. We recommend further application and assessment of the SanQoL measure across other forms of sanitation service and in other settings and populations.

Introduction

Sanitation is the separation of human excreta from human contact (WHO, 2018). Two billion people globally lack access to basic sanitation (UNICEF & WHO, 2019). Improved sanitation can help prevent infectious disease (Cumming et al., 2019; Wolf et al., 2018), but toilet users particularly value other benefits, such as perceived improvements in privacy, safety and dignity (Elmendorf and Buckles, 1980; Jenkins and Curtis, 2005; Solomons, 1978). These are attributes of good quality of life (QoL), and contribute to health in its broadest sense (WHO, 1948). Studies have identified sanitation as a source of psychosocial stress amongst women (Bisung and Elliott, 2017; Kulkarni et al., 2017; Sahoo et al., 2015; Shiras et al., 2018), and found associations between better sanitation and improved mental wellbeing (Caruso et al., 2018).

The sustainable development goal (SDG) indicator for sanitation (6.2.1) measures "safely managed" sanitation services, the definition of which focuses on toilet quality and treatment of waste (UNICEF & WHO, 2019). While this is appropriate for global monitoring, evaluations of specific interventions require a more nuanced understanding of changes brought about in people's lives, such as the extent to which users' perceptions of privacy have been improved. The SDG targets also focus on the user interface and waste treatment, while many other toilet characteristics are important for quality of life, such as whether walls/door provide privacy, whether there is a lock etc. Quantifying incremental improvements in such outcomes is particularly important for economic evaluation methods such as cost-benefit or cost-effectiveness analysis (Drummond et al., 2015). However, economic evaluations of sanitation interventions have always excluded user-reported outcomes beyond disease and time savings as intangible, in the absence of means for their measurement and valuation (Hutton & Chase, 2016).

Given their subjective and psychological nature, aspects of QoL are typically measured as a multi-dimensional psychometric profile or index. For profile measures, a summative scale score with equal weighting is calculated for each dimension, which may or may not be combined into a single overall score (Brazier et al., 2016). For index measures, scores are combined into a single number using a weighting or valuation algorithm. Examples in health economic evaluation are the families of measures denoted "EQ-5D" (Euroqol Group, 2009) and "ICECAP" (Al-Janabi et al., 2012; Grewal et al., 2006). There are two reasons why an index rather than a profile is required for health economic evaluation purposes. First, most applications such as the weighting of years of life or sufficient capability require a single overall score (Drummond et al., 2015; Mitchell et al., 2015). Second, user-derived valuation is required to incorporate trade-offs between dimensions, such that the value of health rather than the state of health is being measured (Brazier et al., 2016). This valuation is usually undertaken using preference elicitation methods such as the time trade-off or discrete choice experiments. In "value sets" thus derived, weighting is not equal per attribute or level, and there can be unequal intervals between levels within an attribute.

The women's sanitation insecurity profile is the only example of a psychometric measure including some QoL attributes of sanitation (Caruso et al., 2017a). It is a 60-item profile measure with seven factor scales, measuring a mixture of experiences, feelings and practices. Its scales have been used as exposure measures to explore associations with mental wellbeing and anxiety (Caruso et al., 2018). However, the measure's lack of an overall score and use of equal weighting preclude its use in economic evaluation. Furthermore, measures used for societal resource allocation would ideally be applicable to the general population rather than women only.

The challenge of measuring and valuing broader QoL in economic evaluations is not unique to sanitation, and applies to public health interventions in general (Weatherly et al., 2009). Health economists are increasingly using capability-based outcome measures as a way to broaden the evaluative space beyond the value of health (Coast et al., 2008c; Greco et al., 2016). The capability approach focuses on the value of what people are able to be and do, rather than the utility (satisfaction) they get from goods, services and activities (Sen, 1993). 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 this study, we aim to develop a measure of sanitation-related quality of life (SanQoL) in urban Maputo, Mozambique, and assess its validity and reliability. Validity is the extent to which an instrument measures what it intends to measure (Fayers and Machin, 2015). Reliability is the extent to which an instrument is free from measurement error. The underlying objective was to enable quantification of the value of sanitation to users in a single score, to support future economic evaluations of sanitation interventions.

Methods and results

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. It was nested within the Maputo Sanitation (MapSan) trial (clinicaltrials.gov registration: NCT02362932), which evalated the impact of a shared urban sanitation intervention on children's enteric infections (Brown et al., 2015). Of Maputo's 1.1 million population (INE, 2019), 70% live in informal settlements (UN-HABITAT, 2010), and approximately 89% use non-sewered sanitation facilities (Hawkins and Muximpua, 2015). More information on the setting, including photos of toilet types, is provided in supplementary material A. This study received approval from the ethics committee of the London School of Hygiene and Tropical Medicine (Ref: 14609) and from the *Comité Nacional de Bioética para a Saúde* (IRB00002657) at the Ministry of Health in Mozambique. All participants provided written informed consent to participate, and all participants were informed of their right to withdraw from the study at any time.

Overall study design

The overall design drew on health-related QoL measures used in economic evaluation, whereby the final measure is an index anchored at 0 and 1 (Drummond et al., 2015). An interim stage, called the "descriptive system", typically comprises a set of psychometric items with categorical response scales (Brazier et al., 2016). Each item represents one dimension of the construct being measured. Therefore, methods such as factor analysis are not required for such measures, and items are selected primarily for content validity, that is, whether the measure represents all relevant aspects of a construct (Fayers and Machin, 2015). In measures of health-related QoL used for economic evaluation, an individual's combination of responses under the descriptive system comprises their "health state" (e.g. 12213). Those states are then valued as an index using trade-off elicitation methods.

We followed this broad approach, aiming to develop a measure of sanitation-related quality of life (SanQoL), using mixed methods. The capability approach 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 with poor sanitation. The primary intended use is in economic evaluation of sanitation programmes. This required a "short-form" approach, with a small number of items, to permit future valuation through trade-off elicitation (Brazier et al., 2016).

We designed a descriptive system of "sanitation states" and evaluated it in the MapSan study population. Initial qualitative research in the same setting (Chapter 5) defined the SanQoL construct as "the aspects of self-perceived quality of life that are directly affected by sanitation practices or services". This draws on common definitions of health-related QoL (Karimi and Brazier, 2016; Peasgood et al., 2014). 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 qualitative research also resulted in a conceptual model (Figure 1) comprising five attributes: health, disgust, privacy, safety and shame.

In this paper, we present the subsequent four stages of measure development (Table 1). Methods and results for each stage are presented concurrently, as the results of each stage informed the design of the next. We adhere to the minimum standards of the International Society for Quality of Life Research (ISOQOL) – the checklist is provided in supplementary material G (Reeve et al., 2013).

Figure 1: Conceptual model for SanQoL

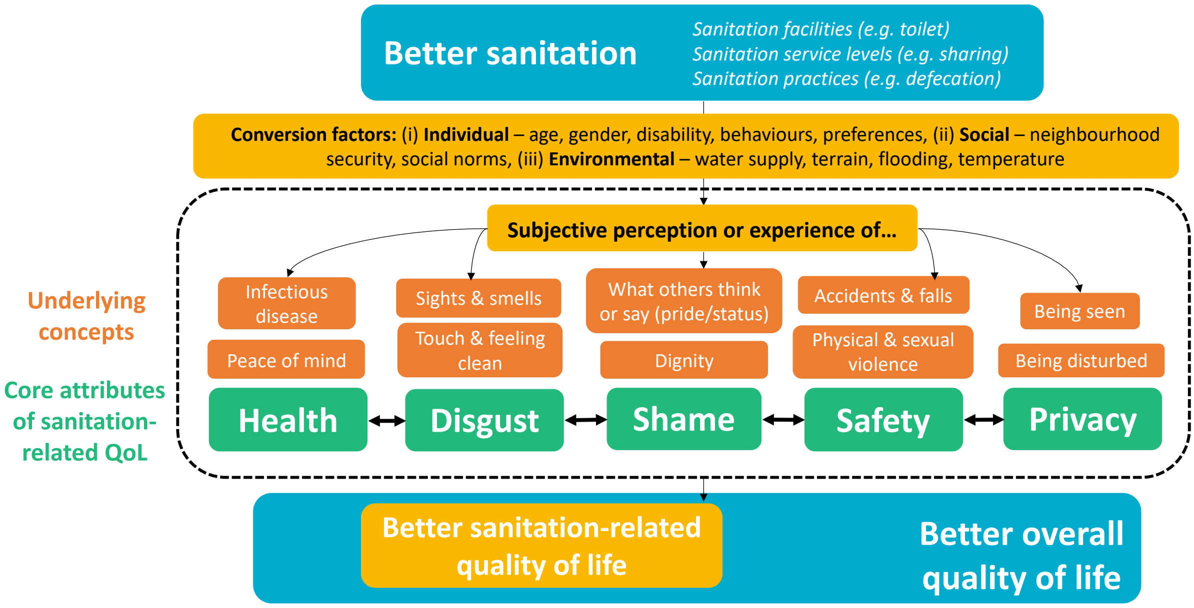


Table 1: Stages followed in measure development after qualitative research

Stage	Objective	Method	Main output
1. Item development & refinement	To develop a list of items for piloting	Long-listing items, review by study team and external experts	Draft descriptive system with one item per attribute
2. Piloting	To ensure items are understood and identify floor/ceiling effects	Pilot interviews (n=64), cognitive interviews (n=28)	Final descriptive system
3. Valuation (weighting) of attributes	To estimating weights for attributes	Rank sum method for visual analogue scale data (n=424)	Weights to calculate SanQoL index values
4. Psychometric testing	To assess validity and reliability	Quantitative survey (n=424)	Evidence of validity and reliability

Stage 1 – Item development

Methods

We developed several possible items for each of the five attributes identified by the qualitative research (Streiner et al., 2015). We considered items phrased with and without capability-oriented wording, e.g. "can you…" versus "how often do you…". To achieve a subjective QoL orientation, items centred on feelings rather than functionings (Fayers and Machin, 2015). We considered response scale formulations based on frequency and intensity, and scales with three, four and five levels. A long-list of 36 items was initially reviewed within the study team and with a further 14 external experts identified by the study team.

Results

For each of the five attributes, we identified one item which appeared best from a face and content validity perspective, considering the qualitative findings. We drafted a five-item descriptive system (Table 2, and in Portuguese in Supplementary Material B) with responses on a four-level ordinal frequency scale: always, sometimes, rarely, never. The descriptive system therefore contains 1,024 sanitation states (4^5). Items were framed as direct questions with capability-oriented wording, such that "always" was the best outcome. We included a "prefer not to answer" option. Four items are phrased as "Can you use the toilet…". The safety item is framed as "Are you able to feel safe while using the toilet?", considered easier to understand than "can you feel safe…". This also avoided a double negative (e.g. "…without feeling unsafe"). Options were extensively discussed in Portuguese as well, before piloting (Stage 2). The preamble for respondents before SanQoL questions was:

> *"Let's talk about your experiences while carrying out your sanitation practices in the past four weeks. By this, I mean any practices you carry out in the toilet you usually use when at home. I am interested in how often you experienced the things in the questions in the past 4 weeks. So, please respond with always, sometimes, rarely or never."*

We specified a four-week recall period to ensure consistency in recall amongst respondents, and to balance the risk of recall bias with the aim of eliciting responses informed by general experience rather than only the last few days (Fayers and Machin, 2015; Kjellsson et al., 2014).

Table 2: Descriptive system for the SanQoL measure

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

*Respondents can choose "prefer not to answer" for any item.

Stage 2 – Piloting

Methods

We undertook piloting and cognitive interviewing during April 2019. We recruited a team of four enumerators (two male and two female), of which two had been part of the qualitative team and one had worked on other sanitation studies in the setting. All spoke fluent Portuguese and Changana, the first and second most commonly-spoken languages in Maputo. ZA, a Mozambican fluent in both languages and with research experience in the study setting, translated items from English into Portuguese. IR, an intermediate Portuguese speaker, extensively discussed meanings with ZA and the enumerator team, and translations of key Changana terms were agreed. We piloted items with 64 individuals from the target population, above the 50 minimum recommended by Streiner et al. (2015) for assessing endorsement frequencies. We undertook further cognitive interviewing with 28 of the pilot respondents, to establish whether items were easily understood and acceptable (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 also rated their explanation on a three-level scale of full comprehension, partial comprehension and no comprehension. Findings were discussed within the team at the end of each day.

Results

Of the 64 piloting respondents, 48% were male and 52% female. Piloting and cognitive interviewing showed that items could be understood. No single category per item received greater than 50% of endorsements, lower than the 80% threshold proposed by Steiner et al. (2015), and there were also no floor or ceiling effects. There were no changes to the descriptive system in English after piloting, but there was one change to the Portuguese. The disgust item (Table 2), originally framed as an adjective (*enojado)* was reframed as a noun *(nojo)*, as this was thought more natural in the Portuguese spoken in the setting.

Stage 3 – Valuation

Methods

Valuation aims to aggregate responses against the descriptive system into a single score, weighted by 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. We aimed to recruit at least 400 people aged over 18 living on MapSan-enrolled compounds, stratified by intervention status and gender. We aimed to recruit two people (one man, one woman) from each compound, on condition that they were not from the same household. This was a suitable study population since a range of SanQoL 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. In addition to SanQoL items, we collected data on water supply and sanitation service usage and demographic characteristics, and calculated a wealth index based on assets included in the most recent demographic and health survey (described in Chapter 7). Following the survey, we asked respondents to rank the five SanQoL attributes using a visual analogue scale (Supplementary Material D).

The vision for SanQoL is to use a "value set" approach based on trade-off elicitation, with an index value for each of the 1,024 sanitation states representing combinations of attribute levels. In the short-term for the present thesis, however, it was not feasible to undertake a subsequent trade-off valuation study with the budget available. We therefore decided to use an interim valuation strategy based on weighting attribute scores. These scores were on a 0-3 scale based on item responses (0 = never, 1 = rarely, 2 = sometimes, 3 = always). We used the rank sum method (Stillwell et al., 1981) to calculate attribute weights from the rank data (Equation 1), following de Kruijk & Rutten (2007) and Greco (2016). We calculated SanQoL index values, which can range from 0-1, based on attribute-level scores and weights (Equation 2). Zero represents no sanitation-related capability and one represents full sanitation-related capability. This builds on the ICECAP measures' framing of anchor points from a philosophical perspective, whereby the best and worst states in the descriptive system are anchored at either end of the 0-1 scale (Coast et al., 2008a).

Equation 1 – attribute weights for a population

$$wt_i = \frac{N - R_i + 1}{\sum_{i=1}^{N}(N - R_i + 1)}$$

Equation 2 – SanQoL index value for an individual

$$S_j = \frac{\sum_{i=1}^{N}(x_{ij} * \mathrm{wt}_i)}{3}$$

where:

wt_i is the weight of the ith attribute

N is the number of attributes

R_i is the mean rank of the ith attribute

x_{ij} are item scores ranging from 0-3 for the jth individual

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 over 60), and treatment group. The rationale for including gender was that women experience sanitation differently to men in that they squat for urination, undertake 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 aging (Groce et al., 2011). In Mozambique, an elderly person is defined in law as anyone aged 60 or older (Castel-Branco & Andrés, 2019). Rank is an ordered categorical variable, so we used mixed-effects ordered logit models, clustering standard errors at the compound level. 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. We conducted all analyses in Stata 16 (StataCorp, 2019).

Results

We sampled 424 individuals (221 female, 203 male) each from a different household (Table 3). They were sampled from 275 MapSan compounds (131 control, 144 intervention) with a response rate of 99%. About two thirds of respondents 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 (82%) of respondents shared their toilet with other households, with a mean of 11.7 people sharing each toilet stance.

Table 3: Characteristics of quantitative sample

	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)	4.7 (3.1)	5.4 (3.0)
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 roof or similar	421 (99%)	203 (100%)	218 (99%)
Household has access to electricity connection	359 (85%)	175 (86%)	184 (84%)
Rents dwelling	114 (27%)	56 (27%)	58 (26%)
Compound-level WASH characteristics			
Piped water connection	416 (98%)	199 (98%)	217 (99%)
Hours/day water available	6.8 (2.3)	6.7 (2.3)	6.8 (2.2)
Uses on-plot toilet	416 (98%)	201 (99%)	215 (98%)
Pour-flush to septic (intervention)	222 (52%)	103 (51%)	119 (54%)
Pit latrine (control)	202 (48%)	100 (49%)	102 (46%)
Shares toilet with other household(s)	349 (82%)	170 (83%)	179 (81%)
Number of households sharing stance	3.1 (1.7)	3.0 (1.7)	3.1 (1.7)
Number of people sharing stance	11.7 (6.0)	11.5 (6.2)	11.9 (5.9)

Data are n (%) for categorical variables and mean (SD) for numerical variables

Health had the highest mean rank and shame the lowest (Table 4), meaning they also had the highest and lowest attribute weights calculated from the above equations. Across the five attributes, there was insufficient evidence at the 5% level for differences in ranks by gender, being elderly, or intervention status in regression models (Table 5). The GLMM models broadly support this result (Supplementary Material C), though suggested that there was good evidence (p=0.037) of elderly people ranking health about a quarter of a rank higher, compared to weaker evidence (p=0.056) for this in the ordered model.

Table 4: Mean ranks of SanQoL attributes and calculated weights

	Mean rank (n=424)						Weight for index valuation (from "overall")
	Overall		Female		Male		
	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

Mean rank data are presented graphically in Supplementary Material C

Table 5: Mixed effects ordered logit (ranks modelled as ordered)

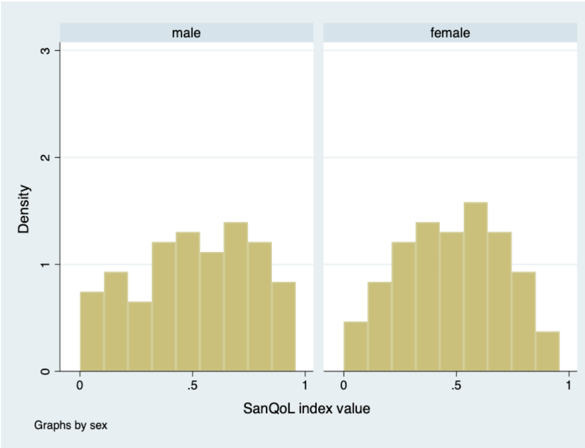
		Disgust	Health	Shame	Safety	Privacy
Female	Odds ratio	0.902	1.065	1.177	1.158	0.787
	p-value	0.561	0.758	0.448	0.437	0.197
Aged 60+	Odds ratio	1.206	0.496*	0.867	1.479	0.763
	p-value	0.492	0.056	0.611	0.166	0.360
Treatment	Odds ratio	0.697*	1.017	1.111	1.342	0.877
	p-value	0.070	0.941	0.677	0.163	0.546

note. coefficients are odds ratios. Standard errors are clustered at the compound level. *, **, *** indicate significance at the 10, 5 and 1 percent level

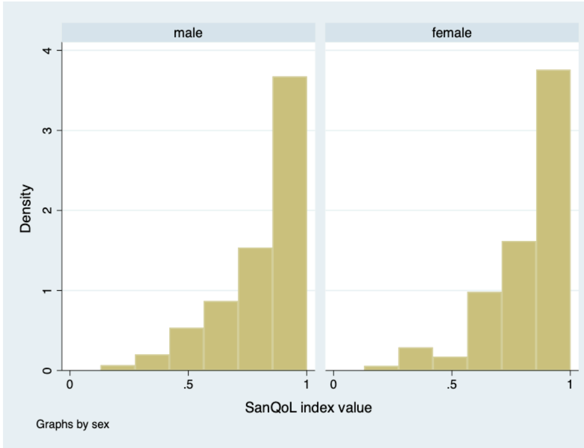
Histograms by gender of SanQoL index values calculated using Equation 2 are presented in Figure 2, for intervention, control and overall. The control group used a more diverse range of toilet types and thus it is notable that histograms for that group approximate normal distributions. In the intervention group, distributions are left-skewed because toilets were high-quality and the modal intervention group respondent achieved the maximum possible level of SanQoL.

Figure 2: Histograms of SanQoL index values by gender and intervention/control

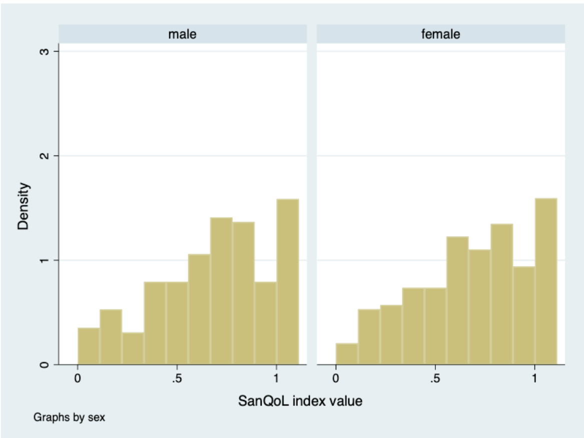
Panel A – control group



Panel B – intervention group



Panel C – overall population



Stage 4 – Validity and reliability assessment

Methods

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 two weeks after the original interview (Streiner et al., 2015). We used a two-way mixed effects model (Koo and

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 about the presence or absence of associations between each attribute score and a set of user and toilet characteristics (Table 6). We tested these using GLMM, because attribute scores represent points on an underlying continuous scale. We regressed each attribute score on all hypothesised covariates, with standard errors clustered at the compound level. The hypotheses drew on previous qualitative work (Chapter 5) and the broader literature on motives for sanitation behaviours and mental wellbeing (Novotný et al., 2018; Sclar et al., 2018). As a robustness check, we analysed item scores as ordered categorical using mixed effects ordered logit models.

We also investigated convergence of SanQoL index values with the WHO-5 mental wellbeing index (Topp et al., 2015). We hypothesised that the correlation coefficient would 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 tested the hypothesis that responses would be correlated but not equal, because any two people may experience the same toilet differently. We assessed this using inter-rater methods, by calculating the ICC using a one-way random effects model (Koo and Li, 2016). Interpretation of this ICC is “fair” (0.40-0.59), “good” (0.60-0.74), or “excellent” (>0.75) (Cicchetti, 1994).

Table 6: Hypothesised presence of absence of associations and rationale

Variable	Type of variable	Association hypothesised					Rationale for hypothesis
		Disgust	Health	Shame	Safety	Privacy	
User characteristics							
Women	Binary	No	No	No	Yes	Yes	Women might have higher acceptability thresholds for safety and privacy, since they are at higher risk of peeping, sexual harassment and assault.
Aged 60+	Binary	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)	Binary	Yes	Yes	Yes	Yes	No	The quality of the toilet floor might affect all attributes, but there is no obvious rationale for privacy except collinearity with other attributes.
Toilet wall material (high/low quality)	Binary	No	No	Yes	Yes	Yes	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	Binary	No	No	No	Yes	Yes	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	Binary	Yes	Yes	Yes	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.

Results

Table 7 reports endorsement frequencies overall and by gender – there were no missing values. 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. 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 0.69-0.73 and Cronbach's alpha was 0.77 (Table 8). For test-retest reliability of the SanQoL index value, the ICC was 0.87 (Table 9). Inter-item polychoric correlation coefficients averaged 0.51, with range 0.40-0.70 (Table 10) supporting the idea that the attributes are sufficiently distinct from one another but nonetheless measure the same construct.

Table 7: Endorsement frequencies for SanQoL items, by gender

Code	Response	Overall (n=424)		Female (n=221)		Male (n=203)	
		N	%	N	%	N	%
1. Disgust							
d1	**Can you use the toilet without feeling disgust?**						
	Always	186	43.9	98	44.3	88	43.3
	Sometimes	125	29.5	64	29.0	61	30.0
	Rarely	28	6.6	15	6.8	13	6.4
	Never	85	20.0	44	19.9	41	20.2
2. Health							
h1	**Can you use the toilet without worrying that it spreads diseases?**						
	Always	174	41.0	96	43.4	78	38.4
	Sometimes	127	30.0	65	29.4	62	30.5
	Rarely	30	7.1	14	6.3	16	7.9
	Never	93	21.9	46	20.8	47	23.2
3. Privacy							
p1	**Can you use the toilet in private, without being seen?**						
	Always	249	58.7	134	60.6	115	56.7
	Sometimes	94	22.2	45	20.4	49	24.1
	Rarely	15	3.5	7	3.2	8	3.9
	Never	66	15.6	35	15.8	31	15.3
4. Shame							
s1	**Can you use the toilet without feeling ashamed for any reason?**						
	Always	198	46.7	105	47.5	93	45.8
	Sometimes	113	26.7	61	27.6	52	25.6
	Rarely	28	6.6	17	7.7	11	5.4
	Never	85	20.0	38	17.2	47	23.2

Code	Response	Overall (n=424)		Female (n=221)		Male (n=203)	
		N	%	N	%	N	%
5. Safety							
t1	**Are you able to feel safe while using the toilet?**						
	Always	193	45.5	90	40.7	103	50.7
	Sometimes	117	27.6	63	28.5	54	26.6
	Rarely	34	8.0	20	9.0	14	6.9
	Never	80	18.9	48	21.7	32	15.8

Table 8: Item-total correlations and Cronbach's alpha

Attribute	observations	item-total correlation	alpha without item
Disgust	424	0.69	0.74
Health	424	0.72	0.73
Privacy	424	0.72	0.72
Shame	424	0.73	0.72
Safety	424	0.73	0.72
	Alpha for whole measure		**0.77**

Table 9: ICCs and kappa for test-retest reliability

	ICC	Kappa (linear-weighted)
Disgust	0.74	0.87
Health	0.64	0.81
Shame	0.65	0.82
Safety	0.83	0.87
Privacy	0.89	0.92
SanQoL index value	0.87	0.87

Table 10: Polychoric inter-item correlations

	Disgust	Health	Shame	Safety	Privacy
Disgust	1.00				
Health	0.56	1.00			
Shame	0.52	0.53	1.00		
Safety	0.41	0.47	0.49	1.00	
Privacy	0.40	0.43	0.54	0.70	1.00

Considering construct validity, Table 11 presents p-values on coefficients for hypothesised associations in GLMM models. Table 12 summarises whether results were in line with these hypotheses – the equivalent for presence hypotheses only is in Supplementary Material C. For 30/35 (86%) hypotheses, the result was in line with the hypothesis considering both presence and absence of an association at the 5% level or lower. Considering only hypotheses for the presence of an association, results were in line for 11/14 (79%). A robustness check using mixed effects ordered logit models produced similar results (28/35 [80%] for all associations and 11/14 [79%] for presence associations) – p-values for these checks are provided in Supplementary Material C.

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.

Table 11: Associations between SanQoL attribute scores and user/toilet characteristics

Variable	Type of variable	P-values on coefficients in GLMM regression on attribute scores				
		Disgust	Health	Shame	Safety	Privacy
User characteristics						
Female	Binary	0.879	0.407	0.320	***<0.001******	0.740
Aged 60+	Binary	0.704	0.943	0.732	0.490	0.558
Wealth index	Continuous	***0.017****	0.278	0.396	0.346	0.950
Toilet characteristics						
Toilet floor material	Binary	0.291	***0.002******	***0.004******	***0.001******	***0.004******
Toilet wall material	Binary	0.140	0.150	0.969	***0.018****	***0.034****
Toilet inside lock	Binary	0.118	0.997	0.083*	***0.006******	***0.002******
Enumerator smells faeces on entry	Binary	***0.005******	***<0.001******	***<0.001******	0.155	0.081*

*, **, *** indicate significance at the 10%, 5% and 1% level, with the 1% level also in bold italic text.

Table 12: Results in agreement with presence or absence hypotheses

Variable	Type of variable	Test results in line with hypothesis?				
		Disgust	Health	Shame	Safety	Privacy
User characteristics						
Women	Binary	yes	yes	yes	yes	no
Aged 60+	Binary	yes	yes	yes	yes	yes
Wealth index	Continuous	no	yes	yes	yes	yes
Toilet characteristics						
Toilet floor material (high/low quality)	Binary	no	yes	yes	yes	no
Toilet wall material (high/low quality)	Binary	yes	yes	no	yes	yes
Toilet inside lock	Binary	yes	yes	yes	yes	yes
Enumerator smells faeces on entry	Binary	yes	yes	yes	yes	yes

Discussion

This study developed and evaluated the “SanQoL” instrument for measuring sanitation-related quality of life. It captures the attributes people most value about sanitation. We followed measure development methods common in health economics for the purpose of economic evaluation. We used qualitative methods to develop a conceptual model organised around five SanQoL attributes: disgust, health, shame, safety and privacy. We identified one psychometric item per attribute to form a five-item descriptive system for measuring sanitation-related QoL. After piloting and cognitive interviewing, we undertook a quantitative survey of the target population to investigate validity and reliability. We valued the index using the rank sum method, based on survey respondents’ rankings of attributes on a visual analogue scale.

Our results provide evidence of the validity and reliability of SanQoL in measuring the value of sanitation to users in low-income urban neighbourhoods of Maputo, Mozambique. The ICCs for test-retest reliability and Cronbach’s alpha were well above commonly-used thresholds. In GLMM modelling we found evidence of association for 86% (30/35) of hypothesised presence or absence associations between the five SanQoL attributes and characteristics of users and toilets. There was convergence with the WHO-5 wellbeing index. The ICC for convergence of SanQoL index values between two respondents using the same toilet indicated substantial but not complete correlation, as hypothesised.

Our measure quantifies sanitation-related QoL in a way which can be valued in economic evaluation, which was not possible previously (Brazier et al., 2016; Hutton and Chase, 2016). It enables quantification of the extent to which users’ perceptions of disgust, health risk, shame, safety and privacy have changed as a result of an intervention. Critically, it then aggregates those perceptions into a single score weighted by relative value of attributes elicited from the target population. This underlies the claim to be measuring the value of sanitation overall, not only the status of diverse attributes. The SanQoL measure can be used in cost-effectiveness analysis (Chapters 8-9), and also in cost-benefit analysis after monetary valuation using willingness to pay methods.

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 measuring individuals' capabilities in relation to their sanitation practices. This is important because sanitation interventions might improve toilets but with heterogeneous and inequitable impacts on QoL, particularly if appropriate attention is not given to gender and social norms (O'Reilly, 2016). The only previous attempt to measure these outcomes was a 60-item profile measure of women's sanitation insecurity (WSI) developed in rural India (Caruso et al., 2017a). However, its lack of an overall score and use of equal weighting mean the WSI profile cannot be used in economic evaluation. Furthermore, the WSI profile is conceptualised as a measure of exposure on the causal pathway to health outcomes (Caruso et al., 2018; Delea et al., 2019), while SanQoL considers the value of sanitation from an intrinsic rather than instrumental perspective (Jain and Subramanian, 2018). Finally, while the WSI profile measures a mixture of experiences, practices and feelings, SanQoL focuses on feelings.

A short-form approach, with one item per attribute, was necessary for permitting future valuation through trade-off elicitation (Brazier et al., 2016). In support of content validity, we note that our SanQoL attributes were commonly identified in previous studies of the household motives for sanitation investment (Novotný 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 (UNICEF & WHO, 2019). 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 most aspects of convenience by including the value of time savings (Hutton and Chase, 2016).

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 in a broader population could be investigated in future research, to establish the effect of response labels on endorsement distributions. However, our distribution (Table 7) aligns with the fact that for the WSI profile measure, also on a four-level frequency scale, only 5% of respondents endorsed the WSI's "often" category in a study with three times the sample size (Caruso et al., 2017a). 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 such a broad response category 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 broader population.

The mean rank data (Table 4) imply that the attributes' order of importance in this population was: health, disgust, privacy, safety, shame. This aligns exactly with results of the pile-sorting exercise undertaken in the earlier qualitative study (Chapter 5). The lack of substantial heterogeneity in rankings by respondent characteristics is supportive of the valuation being broadly appropriate for this population. However, half the sample had received an intervention which contained explicit and implicit health messaging, and the other were likely aware of the intervention in their area. This may have biased their responses in unpredictable ways, but in particular it could have increased the rank given to health due to social desirability bias. Other possible explanations for health being ranked high in this setting (e.g. education, urban setting) were discussed in Chapter 5.

Limitations

Further validation of SanQoL through cognitive interviews and psychometric analysis is recommended for generalisability to other settings with substantially different sanitation practices (e.g. OD, public toilets). For example, respondents practising OD may need questions reformulated in terms of "carrying out your sanitation practices" rather than "using the toilet". In the absence of longitudinal data, we could not assess responsiveness. 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 limit generalisability by providing only a narrow set of sanitation service types and living arrangements – for example, almost all respondents shared their toilet with other households. Further research on the properties of SanQoL would ideally take place in 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 likely experience sanitation in different ways (Sahoo et al., 2015).

The present valuation scheme could be improved upon. Though the rank sum method has previously been used in index valuation (de Kruijk and Rutten, 2007; Greco, 2016), it implicitly assumes equal intervals between response categories (Brazier et al., 2016). While ranking is straightforward for respondents to understand, it encourages trading off of attributes as a whole rather than trading off different levels of attributes, making valuation less nuanced. Future research should apply a trade-off elicitation method, which would address both these concerns (Brazier et al., 2016). We were unable to apply such methods in this thesis, since budget constraints precluded a further survey after the instrument was finalised and validity and reliability investigated as part of the study reported in this chapter.

Achieving safely-managed sanitation for all requires managing faeces along the service chain including transport, treatment and disposal of waste. Our SanQoL measure only captures QoL of toilet users. Additional work would be required to capture QoL of stakeholders at other stages of the chain, for example of sanitation workers involved in emptying pits and septic tanks (World Bank, 2019b). There may also be SanQoL externalities, e.g. one person’s level of sanitation affecting a neighbour’s sanitation-related disgust and perceived health risk.

Conclusion

SanQoL provides a measure of sanitation-related quality of life which enables quantification of the value of sanitation for an individual in a single score, with its five attributes weighted by the target population. It enables measurement and valuation of sanitation attributes which users themselves consider important. Such a measure can support future cost-effectiveness and cost-benefit analyses and provide a more complete accounting of the cost-consequence relationship for sanitation. The SanQoL measure could also be used in impact evaluation or in routine programme 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 sanitation-related quality of life in urban settings of low- and middle-income countries, but further testing in other settings is required.

Funding. 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 interest. None declared.

Acknowledgements. We greatly appreciated the cooperation of survey respondents. We also appreciated the efforts of the fieldworkers employed by WE Consult: 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 and Julia Lohmann. Finally, the inputs of the 14 expert reviewers were helpful in improving items.

Authors' contributions. IR conceived the study design and methodology, as part of a PhD supervised by GG and OC, with RD and CO providing methodological inputs. IR and ZA managed the fieldwork and IR carried out the analyses, with all authors contributing to their interpretation. IR wrote the first draft and collated inputs to all subsequent drafts. All authors reviewed version iterations and contributed intellectual content.

Chapter 7: Research Paper 4 - Evaluation

Chapter introduction

In Chapter 4, I undertook a systematic review which showed substantial weaknesses in the literature on the economic evaluation of sanitation interventions. In particular, there was no means of measuring QoL outcomes. In Chapters 5-6, I developed a measure of sanitation-related quality of life based on the capability approach. This was motivated by wanting to understand the broader benefits of sanitation as valued by users.

The present chapter takes things to the next stage by applying the SanQoL measure. In doing so, I achieve objective 3 of this thesis: to estimate the effect of using better toilets on quality of life, in a quantitative study using the measure. This serves to provide an effectiveness estimate for use in the subsequent cost-effectiveness study (Chapter 9), but the results are also of empirical interest in themselves. The study also provides proof-of-concept that the SanQoL measure can be applied in a trial setting as an outcome, and has useful statistical properties. In this paper, I also develop an alternative measure – a sanitation visual analogue scale.

I am proposing to submit this paper to the journal *World Development* (published by Elsevier) Accordingly, I follow the structure norms in development economics, with integrated results and discussion.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1701671	**Title**	Mr
First Name(s)	Ian		
Surname/Family Name	Ross		
Thesis Title	Measuring and valuing quality of life in the economic evaluation of sanitation interventions		
Primary Supervisor	Giulia Greco		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	World Development
Please list the paper's authors in the intended authorship order:	Ian Ross, Giulia Greco, Zaida Adriano, Rassul Nala, Joe Brown, Robert Dreibelbis, Charles Opondo, Oliver Cumming
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceptualised the study, developed the methods, and coordinated inputs from the supervisory team. I trained the fieldwork team and participated in piloting and data collection. I conducted the analysis and write-up, revising drafts on the basis of co-author feedback.

SECTION E

Student Signature	[redacted]
Date	24/10/20

Supervisor Signature	[redacted]
Date	24/10/20

Do better toilets improve people's lives? Evidence from urban Mozambique

Authors:

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

Affiliations:

[1]London School of Hygiene and Tropical Medicine, UK;

[2]WE Consult, Mozambique;

[3]Instituto Nacional de Saúde, Mozambique;

[4]Georgia University of Technology, USA;

[5]University of Oxford, UK;

Corresponding author:

Ian Ross

LSHTM, 15-17 Tavistock Place, London, WC1H 9SH

Email: ian.ross@lshtm.ac.uk

Phone: 020 7636 8636

Abstract

Systematic review evidence suggests that use of toilets can prevent diarrhoeal disease, but recent trials of sanitation interventions identified little or no effect. Qualitative studies have consistently reported that users value the benefits of toilets for privacy, safety and dignity. However, these "quality of life" outcomes have rarely been measured quantitatively, and never in an evaluation of a sanitation intervention. We took advantage of the existing Maputo Sanitation trial (clinicaltrials.gov, NCT02362932) to evaluate these outcomes in the context of a broader health impact study. We aimed to estimate the effect of a shared toilet intervention on quality of life in urban Mozambique, which comprised a pour-flush interface with concrete superstructure and solid lockable door. Our primary outcome was a novel multi-attribute index of sanitation-related quality of life (SanQoL), combining users' perceptions of disgust, health risk, privacy, safety and shame. Using generalised linear mixed models, we assessed SanQoL index values of people living on intervention and control compounds previously enrolled in the trial. We found strong evidence ($p<0.001$) that the intervention was associated with a 1.6 standard deviation increase in SanQoL as compared to a low-quality pit latrine. This difference was seen across all five attributes of SanQoL, but with particularly large effect sizes for privacy and safety. We identified a similar effect size (1.4 standard deviations) when respondents directly evaluated their level of sanitation using a visual analogue scale. There was also good evidence ($p=0.04$) of a 0.2 standard deviation improvement in mental wellbeing, as measured by the WHO-5 index. These findings demonstrate that better toilets can improve people's lives beyond reducing infectious disease. Our contribution is in quantifying the degree of the difference. Policy-makers may be willing to make decisions on the basis of comparative effectiveness of interventions in quality of life terms. Impact evaluations and economic evaluations could make more use of quality of life measures to inform such decisions.

Keywords: Quality of life, Sanitation, Mozambique, Public health

Introduction

Sanitation is the separation of human excreta from human contact (WHO, 2018). Approximately two billion people globally lack access to a "basic" level of sanitation service (UNICEF & WHO, 2019), and 432,000 annual deaths from diarrhoeal disease are attributable to inadequate sanitation (Prüss-Ustün et al., 2019). Systematic review evidence suggests that toilets prevent diarrhoea (Wolf et al., 2018), but recent intervention trials found little or no effect of sanitation improvements on disease (Cumming et al., 2019).

However, people value sanitation for other reasons than preventing disease. Qualitative studies report benefits of toilets across several domains of quality of life (QoL), including privacy, safety, dignity, pride and cleanliness (Elmendorf and Buckles, 1980; Jenkins and Curtis, 2005). Each of these maps onto one of the seven objective features of QoL identified by a recent commission (Stiglitz et al., 2009) or onto one of the ten central human capabilities identified by Nussbaum (2011). Accordingly, sanitation-related QoL is the subset of overall QoL which is directly affected by sanitation practices or services (Chapter 5). Under the capability approach, value is defined by what people are able to be and do in the pursuit of a good life, rather than utility from consumption (Sen, 1980). In capability terms, sanitation-related QoL is what people are able to be and do with respect to their sanitation practices – for example, whether they are able to defecate in privacy. An intervention which increases people's sanitation-related capabilities improves their QoL.

A household toilet is a private good and sanitation-related QoL is a private benefit (Dickinson et al., 2015). However, an excreta-free environment is a public good in that it is non-rival and non-excludable, and sanitation interventions can have positive public health externalities (Andrés et al., 2017). Autonomous household sanitation investments may therefore comprise private provision of public goods (Vicary, 1997). Studies of individuals' motives for sanitation investment have often identified that pride, privacy, and so on are more important drivers than health (Jenkins and Curtis, 2005; Jenkins and Scott, 2007). An individual's expected QoL payoff from a proposed private sanitation investment is therefore an important determinant of whether the public good of an excreta-free environment is achieved. Different sanitation intervention options proposed for public investment will succeed in improving sanitation-related QoL to different extents. It is surprising, then, that aspects of sanitation-related QoL have rarely been measured quantitatively, and never assessed in impact evaluations.

We know of only two quantitative studies which explored the association of improvements in sanitation with aspects of QoL. A cross-sectional study in rural India found that latrine access was associated with higher scores on the WHO-5 mental well-being index (Caruso et al., 2018). Another cross-sectional study found a similar positive association with WHO-5 scores in urban Bangladesh (Gruebner et al., 2012). We are not aware of studies which quantified the QoL effect of a specific intervention.

In this study, we aim to evaluate the effect of a shared sanitation intervention on quality of life and mental wellbeing. We do so in the context of low-income urban areas in Maputo, Mozambique's capital city. Our study is nested within the broader Maputo Sanitation (MapSan) trial (clinicaltrials.gov, NCT02362932), a health impact study of a shared sanitation intervention (Knee et al., 2020).

Methods

Context

Maputo has a population of 1.1 million (INE, 2019) of which 41% use pit latrines (Hawkins and Muximpua, 2015). Our study site comprises low-income neighbourhoods in an approximately 10km^2 area of the Nhlamankulu district (map in Supplementary Material A). In this area, housing is diverse but the poorest people live in informally-walled 'compounds', where many families share the same toilet and courtyard space (Brown et al., 2015). Houses in such compounds are typically single-storey one-room dwellings. Low-quality pit latrines are common, with unlined pits and squatting slabs made of wood, tyres or concrete, and no water seal (u-bend) providing a barrier to 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 (images in supplementary material A).

Study design

This is an observational study, in which we capitalise on an existing health impact evaluation to assess differences in outcomes between two cohorts. Our sample is drawn randomly from people residing in compounds in the intervention and control groups of an earlier non-randomised trial. The MapSan trial evaluated the impact of a sanitation intervention on childhood enteric infections, employing a controlled before and after design (Knee et al., 2020). As each MapSan intervention compound was identified, a control compound was concurrently matched by cluster size and intervention siting criteria, and enrolled (Brown et al., 2015).

The intervention was implemented by Water and Sanitation for the Urban Poor (WSUP), an international non-government organisation (NGO). The intervention provided a subsidised toilet with a pour-flush toilet user interface with water seal, concrete superstructure, and lockable door. The interface discharges to a septic tank with soakaway. There were two main design types depending on user numbers. The first was a shared toilet (ST) with one stance (cubicle) to be used by around 15 people, at 85% subsidy. The second was a Community Sanitation Block (CSB) with two stances, to be used by a minimum of 21 people, at 90% subsidy. Both STs and CSBs are built from concrete blocks, with metal doors lockable from the inside. Compounds were eligible to receive the intervention if households were sharing a low-quality pit latrine and willing to pay the 10-15% capital contribution – approximately US$ 120 for CSB (2015 prices) and US$ 80 for ST (Mattson, 2016). Further details of the setting and intervention are provided in Supplementary Material A and the trial protocol (Brown et al., 2015). Many aspects of the overall intervention design contribute to privacy and safety without any likely infectious disease consequences, such as the solid walls and lockable door. In what follows, I use “toilet” to refer to the whole infrastructure, not only the user interface.

Participants

Eligible participants were defined as people aged 18 or over who: (i) were living on intervention or control compounds enrolled in MapSan; (ii) had been living there for at least four years, since before the intervention; (iii) were using the type of toilet consistent with intervention or control status of the compound (e.g. still using a pit latrine if control). The first and second criteria ensured that, prior to the intervention, all our participants had been using a pit latrine without a water seal on that same compound they still lived on. This aimed to reduce risk of selection bias, because there has been migration out of and into MapSan-enrolled compounds since 2015 (Shiras et al., 2018), and the characteristics of joiners or leavers may have been influenced by better-quality sanitation facilities. The second criterion also ensured that intervention group participants had experience using poor-quality toilets in this setting, noting Sen's (1999) concerns about adaptation to poor circumstances hampering interpersonal comparisons of wellbeing. One reason for the third criterion was the *ex-ante* knowledge that some control compounds had received NGO interventions after the MapSan endline, and others had autonomously upgraded their own toilets.

Outcomes

We measured quality of life in three ways. The primary outcome was an index of sanitation-related quality of life (SanQoL), a multi-attribute measure based on qualitative research in this setting (Chapter 6). SanQoL measures the aspects of self-perceived quality of life which are directly affected by sanitation practices or services, with psychometric items based on the capability approach (Sen, 1980). The scope of "sanitation practices" is as determined by the user, but is assumed to include defecation, urination, bathing and menstrual hygiene (Chapter 5). Validity and reliability of the measure was assessed through piloting, cognitive interviews, and psychometric analysis (Chapter 6).

The five SanQoL attributes are disgust, health, privacy, safety and shame. Each is assessed by a single psychometric item (Table 1), measured on a four-level frequency response scale: always (3), sometimes (2), rarely (1), never (0). Responses are aggregated into an index by weighting attributes according to the population's opinion of their relative importance. To estimate weights, respondents were asked to rank the five SanQoL attributes using laminated labels on a velcro board (Supplementary Material B).

Table 1: SanQoL dimensions, questions and response categories

#	Attribute	Psychometric item	Responses	Weight in index values
1	Disgust	Can you use the toilet without feeling disgusted?	3 - Always 2 - Sometimes 1 - Rarely 0 - Never	0.22
2	Health	Can you use the toilet without worrying that it spreads diseases?		0.29
3	Privacy	Can you use the toilet in private, without being seen?		0.20
4	Shame	Can you use the toilet without feeling ashamed for any reason?		0.13
5	Safety	Are you able to feel safe while using the toilet?		0.16

Using data on mean ranks, we estimated weights (Equation A) for each SanQoL dimension using the rank sum method (Stillwell et al., 1981), as previously used in index valuation by de Kruijk & Rutten (2007) and Greco (2016). The ensuing weights were used to calculate SanQoL index values on a 0-1 scale (Equation B). Zero represents 'no sanitation-related capability' and one 'full sanitation-related capability', drawing on earlier capability-based measures (Coast et al., 2008a).

Equation A – attribute weights for a population

$$wt_i = \frac{N - R_i + 1}{\sum_{i=1}^{N}(N - R_i + 1)}$$

Equation B – SanQoL index value for an individual

$$S_j = \frac{\sum_{i=1}^{N}(x_{ij} * \mathrm{wt}_i)}{3}$$

where:

wt_i is the weight of the ith attribute

N is the number of attributes

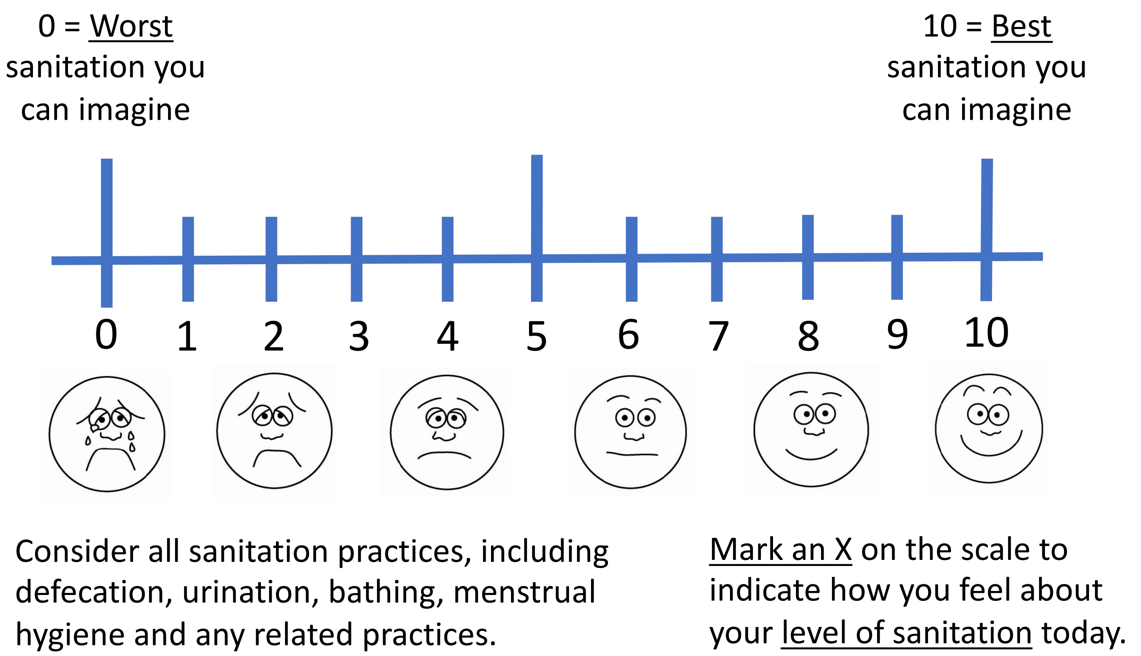
R_i is the mean rank of the ith attribute in the population

x_{ij} are item scores ranging from 0-3 for the jth individual

S_j is the SanQoL index value for the jth individual

Two secondary outcomes were considered. The first was a sanitation visual analogue scale (VAS). We asked people to indicate on a paper-based 0-10 scale how they felt about their "level of sanitation today", where zero is “worst imaginable sanitation” and ten is “best imaginable sanitation” (Figure 1). This was adapted from the VAS in the EQ-5D measure of health-related quality of life (Euroqol Group, 2009) with emoji visualisation informed by the visual pain scale (Hawker et al., 2011). The enumerator read out the guidance (Figure 1), then the respondent indicated their selected level on the scale with a pencil. Second, we used the WHO-5 mental wellbeing index, a widely-used multi-attribute instrument for assessing subjective mental well-being (Topp et al., 2015). It comprises five questionnaire items rated on a frequency scale, related to feeling cheerful, calm, active, well-rested, and finding enjoyment in daily life. Scores are summed and rescaled to 0-100, with a higher score interpreted as better mental well-being.

Figure 1: Sanitation visual analogue scale (VAS)

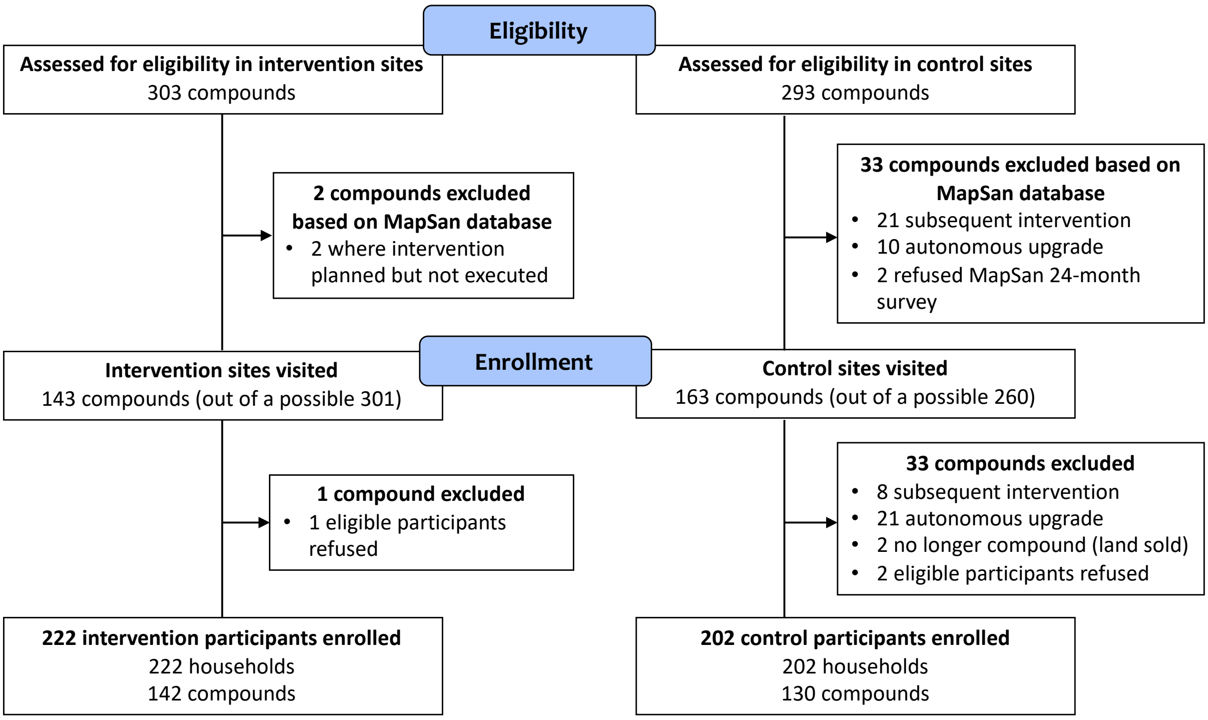


The rationale for these three outcomes was as follows. SanQoL is a sanitation-specific measure of QoL developed and validated in this setting. The sanitation VAS was included to explore whether a comparable effect (and effect size) would be seen when people rated their level of sanitation on a scale directly, as opposed to indirectly via a multi-attribute instrument with population-based weights. The VAS takes less time to complete than the SanQoL questions, so may be easier to include future studies if it provides a suitable proxy. It may also be used when appropriate value sets for converting SanQoL item responses to index values are not available. We included the WHO-5 because mental wellbeing is one aspect of broader quality of life thought to be influenced by sanitation (Sclar et al., 2018), and the WHO-5 has previously been used in sanitation studies (Caruso et al., 2018; Delea et al., 2019). All three outcomes measure aspects of QoL. However, while the type of sanitation service a person uses may be an important determinant of their SanQoL and sanitation VAS scores, it is likely to be only a minor factor in affecting mental wellbeing.

Sampling

There were only 561 eligible MapSan compounds before encountering unknown levels of emigration and upgrading during data collection (Figure 2). We aimed to recruit two people per compound (one man, one woman) from different households, which introduced clustering at the compound level. The sample size calculation was estimated according to a formula for the comparison of two means with power at 80% and significance at 0.05. The required sample size to detect a 0.05 mean difference in SanQoL with a standard deviation of 0.15 and intracluster correlation coefficient (ICC) of 0.4 would be 398. We undertook face-to-face interviewing using smartphones during April-June 2019, with the mWater (2019) surveyor application. The list of MapSan intervention and control compounds was randomised and lists allocated to enumerators. Upon arrival at each compound on their list and with the approval of a resident, fieldworkers inspected the toilet for the presence of a water seal (u-bend). Next, by talking to a resident, they listed all eligible people based on the inclusion criteria. Sampling was random from the list of eligible men within the compound, by approaching households starting from the house opposite the compound entrance, and working leftwards until an eligible man was identified. The same process was followed for eligible women, with the condition that the female respondent not be from the same household as the interviewed man. We continued visiting compounds until the target sample size was reached. Interviews were in Portuguese, unless the respondent preferred to talk in Changana, a local language in which all interviewers were fluent.

Figure 2: Stages in eligibility and enrolment



Statistical analysis

We analysed participants according to trial arm, in order to test the overarching hypothesis that the intervention was associated with an improvement in quality of life. We tested specific hypotheses about the association of the intervention with SanQoL index values, VAS scores and mental wellbeing (Table 2), and interactions with gender and being elderly. We also investigated the relative size of effects at the level of individual SanQoL attributes, and assessed SanQoL effects for ST users and CSB users separately.

To test hypothesis 1a (Table 2) we used a generalised linear mixed model (GLMM), with gaussian distribution and identity link, and a random effect at the compound level to account for clustering:

$$S_{ij} = \alpha_0 + \alpha_1 T + \boldsymbol{\beta X}_{ij} + u_i + \varepsilon_{ij}$$

where:

S_{ij} represents the SanQoL index value for individual j in compound i

T is a binary variable set at 1 for intervention toilet users and 0 for PL users

$\boldsymbol{X}_{ij}$ is a vector of covariates

α_0 is a constant which has no interpretation in this case

α_1 is a coefficient and $\boldsymbol{\beta}$ a vector of coefficients

u_i is a random effect at the compound level

ε_{ij} is the error term

Standard errors were clustered at the compound level, since the intervention was applied at this level. We included two types of covariates in $\boldsymbol{X}_{ij}$. First, we adjusted for characteristics which were unbalanced at the 5% level between groups (Table 3). Second, in all adjusted models we included binary variables for gender and being elderly in $\boldsymbol{X}_{ij}$, since these were hypothesised to affect the association between exposure and outcome (Table 2). For hypotheses 1b and 1c, we estimated a GLMM with the same covariates but a different dependent variable as appropriate. For hypotheses 2a and 2b, we specified the same models but including a factorial interaction with T for the gender or elderly binary variables respectively. However, our sample was not powered for sub-group analyses. We analysed data in Stata 16 (StataCorp, 2019). We computed a wealth index using principal components analysis following standard practice (Vyas and Kumaranayake, 2006), using the asset list from the most recent Demographic and Health Survey in Mozambique (MISAU et al., 2013).

Table 2: Hypotheses

	Hypothesis	Rationale
1. Main	(a) the intervention is associated with higher SanQoL index values	The intervention should improve people's sanitation-related capabilities and self-perceived level of sanitation in general (Chapter 5; Shiras et al. 2018)
	(b) the intervention is associated with higher sanitation VAS score	
	(c) the intervention is associated with higher mental wellbeing (WHO-5)	Sanitation interventions may improve people's self-perceived mental wellbeing, as shown in a systematic review of mainly qualitative studies (Sclar et al., 2018) and cross-sectional studies using WHO-5 (Caruso et al., 2018; Gruebner et al. 2012).
2. Sub-groups	(a) across all three outcomes, any intervention effect is larger for women than men	Women experience sanitation differently to men in that they squat for urination, undertake menstrual hygiene, and are more likely to fear and experience assault (Tilley et al., 2013). Sanitation deficits particularly impact on women (O'Reilly, 2016), so improvements in sanitation may disproportionately benefit women more.
	(b) across all three outcomes, any intervention effect is larger for elderly people (aged 60+) than non-elderly	Elderly people are impacted by poor sanitation, predominantly as a result of disabilities that occur with aging (Groce et al., 2011). Therefore, improvements in sanitation may disproportionately benefit elderly people. In Mozambique, an elderly person is defined in law as anyone aged 60 years or older (Castel-Branco and Andrés, 2019).

Robustness checks

We assessed the sensitivity of results to different modelling approaches (Lu and White, 2014). First, we included in $\boldsymbol{X}_{ij}$ only covariates significantly different between groups at the 10% level (Table 3) and excluded the gender and elderly binary variables. Second, we instead included all covariates hypothesised *ex ante* as influencing sanitation-related QoL, to explore the presence of omitted variable bias (gender, elderly, renting dwelling, number of people sharing a toilet stance, whether the toilet is shared with other households). Third, to explore omitted variable bias in the WHO-5 regressions, we included covariates hypothesised to affect mental wellbeing (gender, being elderly, whether the respondent has a partner, is in pain, or has problems walking). Fourth, we explored whether using a generalised estimating equations (GEE) or ordinary least squares (OLS) specification instead of GLMM affected the main results, again with standard errors clustered at the compound level. Finally, we explored whether using mixed-effects ordered logit for the SanQoL attribute-level regressions, rather than treating them as continuous variables, made a difference to results.

Ethics

Our study received prior approval from the Research Ethics Committee of the London School of Hygiene and Tropical Medicine (Ref: 14609) and from the *Comité Nacional de Bioética para a Saúde* (IRB00002657) at the Ministry of Health in Mozambique. Informed, written consent was obtained from all participants.

Results and discussion

Participant characteristics

We sampled individuals from 424 different households across 272 compounds (clusters), of which 130 were control and 142 intervention (Figure 2). In some compounds, only one man or woman was eligible and available (mean number of respondents per cluster: 1.6). The response rate amongst eligible respondents approached was 99%. There was insufficient evidence of difference in characteristics between intervention and control groups at the 5% level, except for the wealth index score (Table 3). People living on intervention compounds were slightly wealthier than controls, but assets with the most different levels of ownership were the less expensive ones (e.g. watch, radio), while more expensive ones (e.g. concrete walls, fridge, television) were not significantly different. Nonetheless, wealth might be associated with unobserved confounders of the effect of using a better-quality toilet on SanQoL. We anticipated that almost all respondents would be sharing a toilet with other households. In the event, this was the case for 81% of control and 83% of intervention households (Table 3). These were single-household compounds, likely due to empty dwellings (driven by rental markets or migration) or changes in compound living arrangements in the four years since the intervention.

Table 3: Characteristics of sample

	Control (n=202)	Intervention (n=222)	P-value for difference
Demographic characteristics			
Respondent is male	101 (50%)	103 (46%)	0.459
Respondent age	38.4 (14.9)	41.2 (15.6)	0.059*
Respondent has a partner	107 (53%)	107 (48%)	0.327
Household size	4.9 (2.8)	5.2 (3.2)	0.323
Number of children under-14	1.4 (1.5)	1.2 (1.6)	0.122
Wealth index			
Wealth index score	-0.13 (1.00)	0.12 (0.99)	0.010**
Dwelling has cement or tiled floor	*184 (91%)*	*210 (95%)*	*0.160*
Dwelling has concrete exterior walls	*140 (69%)*	*143 (64%)*	*0.287*
Access to electricity connection	*167 (83%)*	*192 (86%)*	*0.277*
Access to piped water connection	*199 (99%)*	*217 (98%)*	*0.563*
Household cooks indoors	*114 (56%)*	*114 (51%)*	*0.295*
Household owns television	*153 (76%)*	*184 (83%)*	*0.069**
Household owns fridge	*98 (49%)*	*128 (58%)*	*0.060**
Household owns mobile phone	*166 (82%)*	*191 (86%)*	*0.278*
Household owns bicycle	*7 (3%)*	*6 (3%)*	*0.650*
Household owns radio	*63 (31%)*	*96 (43%)*	*0.010***
Household owns watch	*89 (44%)*	*130 (59%)*	*0.002****
Other respondent characteristics			
Respondent completed primary school or above	128 (63%)	140 (63%)	0.949
Respondent completed secondary school or above	18 (9%)	33 (15%)	0.060*
Respondent has moderate problems walking about, or worse	12 (6%)	13 (6%)	0.971
Respondent has moderate pain or discomfort, or worse	21 (10%)	17 (8%)	0.325
Respondent rents dwelling	60 (30%)	54 (24%)	0.213
Respondent's dwelling has zinc roof or similar	202 (100%)	222 (100%)	n/a
Compound-level WASH characteristics			
Hours/day water available	6.7 (2.4)	6.9 (2.2)	0.276
Uses on-plot toilet	197 (98%)	219 (99%)	0.397
Shares toilet with other household(s)	164 (81%)	185 (83%)	0.564
Number of households sharing stance	3.1 (1.8)	3.0 (1.6)	0.626
Number of people sharing stance	11.2 (5.5)	12.2 (6.5)	0.078*

Data are n (%) for categorical variables and mean (SD) for numerical variables. *, **, *** indicate significance at the 10, 5 and 1 percent level. Variables included in the wealth index are italicised below the wealth index score.

Was the intervention associated with improved sanitation-related quality of life?

The adjusted model (Table 4, row 1) provides strong evidence ($p<0.001$) that the intervention was associated with a difference in SanQoL of 0.34 (95% CI: 0.29-0.39). Recalling that SanQoL is on a 0-1 scale, the effect size was 1.6 standard deviations (SD), interpreted as 'very large' (Cohen, 1988). This result supports hypothesis 1a that the intervention was associated with better sanitation-related quality of life. Full regression results for all models are in Supplementary Material D, including diagnostic plots of residuals.

To put this into context, mean SanQoL index value amongst control compound respondents was 0.49 (95% CI: 0.46-0.53), compared to 0.83 (95% CI: 0.81-0.86) amongst intervention compound respondents (Table 4). The distribution of SanQoL amongst control compound respondents was broad (Supplementary Material C), since toilet quality was diverse. The modal index value in the intervention group was 1, suggestive of a ceiling effect. The SanQoL difference was larger for users of STs at 0.35 (95% CI: 0.30-0.40) than CSBs at 0.28 (95% CI: 0.20 – 0.36) though confidence intervals overlap (Supplementary Material D).

The intervention was associated with 2.9 additional points on the 0-10 VAS (95% CI: 2.4 - 3.4) in the adjusted model (Table 4, row 2). This is analogous to a difference in SanQoL index value of 0.29, and is supportive of hypothesis 1b. A similar effect size (1.3 SD) to SanQoL (1.6 SD) is seen when the value of sanitation is assessed directly by users on a VAS. This comes despite the fact that VAS scores are distributed quite differently to SanQoL (Supplementary Material C). For example, VAS scores in the control group approximate a normal distribution with a set of people at 0, while SanQoL scores are more spread across the scoring range and unimodal. In the intervention group the modal VAS score was 8/10 while in SanQoL it was 1.0.

Solid walls and doors were likely to have improved perceptions of privacy, safety and shame. The pour-flush mechanism in intervention toilets was likely to have reduced smells and visible faeces compared to pit latrines, thereby possibly improving perceptions of disgust, shame, and health risk. A discrete choice experiment in urban Zambia illustrates that such toilet characteristics are valuable to people, by exploring renters' willingness to pay (WTP) for toilet attributes (Tidwell et al., 2019). The authors found that renters' stated WTP additional rent was US$ 3.4 per month (8% of the median rent of US$ 45) for solid toilet doors and about US$ 2.2 more (5%) for flush toilets as opposed to pit latrines.

For the broader purpose of this study in facilitating economic evaluation, what is important is that the size of the difference in SanQoL has been quantified, allowing cost-effectiveness to be assessed. The outcome being measured is specific to sanitation, noting the questions in Table 1. This approach is common in health economics, with outcome measures for every conceivable disease area or intervention type (Brazier et al., 2016). Whether the measure captures QoL domains affected by asthma, arthritis or acne, only a small subset of domains of overall QoL are affected. Tailored QoL measures such as SanQoL capture only what people value about that area of life.

Was the intervention associated with improved mental wellbeing?

There was good evidence (p=0.041) for the intervention being associated with improved mental wellbeing (Table 4, row 3). The adjusted difference was 6.2 points (95% CI: 0.3-12.2), recalling that WHO-5 is on a 0-100 scale. The effect size was 'small' at 0.2 SD (Cohen, 1988), which is not surprising since sanitation is only one of many factors affecting mental wellbeing (Sclar et al., 2018). Two previous cross-sectional studies identified associations of sanitation improvements with WHO-5, providing a precedent for our result (Caruso et al., 2018; Gruebner et al., 2012). The novelty in our study is in evaluating a specific service level transition as a result of an intervention, rather than assessing many types of services in cross-section.

Table 4: Effects on primary and secondary outcomes

		Means		Unadjusted models		Adjusted models		
	Outcome	**Control (n=202) Mean (SE)**	**Intervention (n=222) Mean (SE)**	**Unadjusted difference (95% CI)**	**p-value**	**Adjusted difference (95% CI)**	**p-value**	**Adjusted effect size (Cohen's *d*)**
1	**SanQoL (0-1 scale)**	0.49 (0.02)	0.83 (0.01)	0.34*** (0.29 - 0.38)	<0.001	0.34*** (0.29 - 0.39)	<0.001	1.6
2	**Sanitation VAS (0-10 scale)**	4.1 (0.2)	7.0 (0.1)	2.9*** (2.4 - 3.4)	<0.001	2.9*** (2.4 - 3.4)	<0.001	1.3
3	**WHO-5 (0-100 scale)**	54.4 (1.9)	58.7 (1.9)	5.6* (-0.4 - 11.6)	0.065	6.2** (0.3 - 12.2)	0.041	0.2

Adjusted models include gender, aged over 60, and wealth score as covariates. Standard errors are clustered at the compound level. *, **, *** indicate significance at the 10, 5 and 1 percent level. Detailed regression output is in Supplementary Material D.

Table 5: Interactions by sex and elderly

		Gender interaction model	**Age interaction model**
	Outcome	p-value on coefficient for Female * Intervention interaction	p-value on coefficient for Aged 60+ * Intervention interaction
1	SanQoL	0.49	0.62
2	Sanitation VAS	0.37	0.98
3	WHO-5	0.84	0.47

Models include gender, aged over 60, and wealth score as covariates. Standard errors are clustered at the compound level. Interactions were coded as full factorial (## in Stata). Detailed regression output is in Supplementary Material D.

Table 6: Differences in individual SanQoL attributes and interactions with gender and age

		Outcome: SanQoL attribute scores (ranging from 0-3)				
		(1) Disgust	**(2) Health**	**(3) Shame**	**(4) Safety**	**(5) Privacy**
Means by group	Control (n=202) Mean (SE)	1.59 (0.082)	1.40 (0.085)	1.56 (0.081)	1.29 (0.080)	1.58 (0.081)
	Intervention (n=222) Mean (SE)	2.32 (0.067)	2.36 (0.060)	2.40 (0.068)	2.64 (0.044)	2.84 (0.037)
Main model	Unadjusted difference (95% CI)	0.72 (0.50 - 0.94)	0.96 (0.74 - 1.18)	0.82 (0.6 - 1.04)	1.35 (1.16 - 1.54)	1.26 (1.07 - 1.45)
	p-value	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***
	Adjusted difference (95% CI)	0.75 (0.53 - 0.97)	0.96 (0.74 - 1.18)	0.80 (0.58 - 1.02)	1.36 (1.16 - 1.56)	1.25 (1.06 - 1.44)
	p-value	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***
	Adjusted effect size (Cohen's d)	0.7	0.9	0.7	1.5	1.4
Gender interaction model	p-value on coefficient for Female*Intervention interaction term	0.56	0.98	0.19	0.29	0.83
Age interaction model	p-value on coefficient for Age*Intervention interaction term	0.45	0.25	0.70	0.07†	0.78

Models include gender, aged over 60, and wealth score as covariates. Standard errors are clustered at the compound level. Interactions were coded as full factorial (## in Stata). *, **, *** indicate significance at the 10, 5 and 1 percent level. Detailed regression output is in Supplementary Material D. †coefficients imply that the intervention's effect on the safety score amongst people under 60 was 1.42 points and on people 60+ it was 0.93 points (supplementary material D).

Were intervention effects larger for women and the elderly?

For all three outcomes, there was insufficient evidence that women benefitted more from better toilets than men, or elderly people more than non-elderly (Table 5). Our study was not powered for these analyses, making the results exploratory. While the sample was gender-balanced (Table 3), there were only 55 participants aged 60 or over, compared to 369 aged under 60. Nonetheless, the p-values on interaction terms (Table 5) are large for all sub-group analyses. In the SanQoL regressions without interaction terms (Table 4), neither the gender nor aged 60+ covariates were significant at the 5% level (supplementary material D). In the VAS regressions, however, there was weak evidence (p=0.056) for women having slightly lower (-0.3) VAS scores than men across the sample as a whole (Supplementary Material D).

In the urban Zambia WTP study discussed above (Tidwell et al., 2019), the authors found insufficient evidence that women's WTP was any higher or lower than men's for flushing toilets or solid doors, either in discrete choice or contingent valuation data (Tidwell, personal communication, 29th April 2020). A revealed preference WTP study in Tanzania found that when the respondent was male, subsequent redemption of vouchers for latrine slabs was 29% higher than for women (Peletz et al., 2017). A stated preference WTP study in Kenya found insufficient evidence for a gender difference in WTP for latrine slabs (Peletz et al., 2019). There is plenty of qualitative evidence for gendered experience of sanitation (O'Reilly, 2016; Tilley et al., 2013). There is less qualitative evidence specific to elderly people, because the majority of their sanitation challenges are a consequence of intersection with disability, rather than simply as a result of being old (Groce et al., 2011; Wilbur and Gosling, 2018). However, from the perspective of interventions, we are not aware of any quantitative evidence for sanitation interventions disproportionately benefitting women or older people, for any outcome.

Effects at the SanQoL attribute level

There was strong evidence for an association of the intervention with differences in all SanQoL attributes, but the size of difference varies (Table 6). Larger effect sizes were seen for safety (1.5 SD) and privacy (1.4 SD), driven by higher scores than other attributes in the intervention group, as scores were relatively similar across attributes in the control group (Figure 3). There was insufficient evidence of any interactions of the intervention with gender for any attribute (Table 6). For safety, there was good evidence (p<0.001) that in the sample as a whole women felt slightly less safe than men (-0.3 on a 0-3 scale) (Supplementary Material D). Women's perception of safety was particularly low on control compounds, and women still felt slightly less safe than men in the intervention group (Figure 3). For age, there was weak evidence (p=0.07) that, while the intervention improved perceptions of safety for elderly people, the improvement was 35% smaller than for non-elderly (Supplementary Material D).

Figure 3: Gender differences in SanQoL attribute scores (male n = 202, female n = 222)

Distributions of SanQoL index values and attribute-level scores by intervention and control (Supplementary Material C) show that responses of people on control compounds were more heterogeneous than on intervention compounds. This reflects the fact that the control toilets themselves were more heterogeneous in quality than the intervention ones, which were constructed to the same two designs. However, SanQoL attribute-level scores also illustrate ceiling effects in the intervention group. For all five attributes, the modal category was the highest possible ("always"). Inter-item polychoric correlation (Supplementary

Material C) shows that privacy and safety were most correlated (0.70) and disgust and safety least correlated (0.41).

Robustness checks

First, including only covariates significantly different between groups at the 10% level (Table 3) made no difference to results for any of the three outcomes (Supplementary Material E). Second, when all covariates hypothesised *ex ante* as influencing SanQoL were included, there was no evidence of omitted variable bias in terms of the size and p-value of the coefficient on the intervention variable (Supplementary Material E). However, the binary covariate for sharing the toilet with other households was significant at the 1% level, which is further explored as a factorial interaction in Supplementary Material E. Amongst the sub-group sharing toilets with other households, there was strong evidence ($p<0.001$) that the intervention was associated with a difference in SanQoL of 0.38 (95% CI: 0.33 – 0.43). This is greater than the difference of 0.34 in the sample as a whole (Table 4). By contrast, for VAS scores, the difference was 2.7 (95% CI: 2.1 – 3.2), smaller than the difference of 2.9 in the sample as a whole. For WHO-5, there was no longer any evidence of a difference amongst the sub-group using shared toilets, though amongst users of private toilets there was a substantial difference of 19.9 (95% CI: 6.6 – 33.2), compared to 6.2 in the sample as a whole (Table 4). It is important that only four years after the intervention, the benefits of toilets which were meant to be shared were in fact being enjoyed by only one household (with mean size 6) in 20% of intervention compounds. Further results and discussion of sharing are in Supplementary Material E.

Third, when all covariates hypothesised *ex ante* as influencing mental wellbeing were included in the WHO-5 regression, there was no evidence of omitted variable bias in terms of the size and p-value of the coefficient on the intervention variable (Supplementary Material E). This means no change to our result with respect to the hypothesis. However, the covariates for pain and problems walking were significant at the 1% and 5% level respectively, and the size of the coefficient on the aged 60+ covariate reduced by half. Fourth, using a GEE or OLS specification did not affect headline results for SanQoL or VAS (Supplementary Material E). In the WHO-5 OLS regression, there was no longer sufficient evidence for an intervention effect, but residuals were bimodally distributed and standard errors far larger than in GLMM and GEE models, indicating OLS is unlikely to be appropriate for this data.

Finally, using mixed-effects ordered logit for SanQoL attribute-level regressions, rather than treating them as continuous variables in GLMM, made no difference to interpretation. The coefficients on the intervention variable remained significant at the 1% level for all SanQoL attributes (Supplementary Material E), and the negative safety effect for women is again observed. Likewise, the same results were seen in the gender and elderly interaction models using ordered logit, including the weak evidence for an interaction with safety for people aged 60+. Overall, we conclude from these robustness checks that our main models (Table 4) were appropriate for testing our hypotheses.

Limitations

Our study has several limitations. It relied on the 'controlled before and after' design of the MapSan trial (Brown et al., 2015), which has inherent limitations. In particular, this design precludes randomisation which risks selection bias in the allocation of intervention and control groups. It also necessitates adjusting for covariates, but these may be imprecisely measured or unmeasured, making it impossible to eliminate the risk of unobserved confounding. We relied on MapSan allocation groups in our enrolment strategy and, with SanQoL being a novel measure, it was only possible to collect data after the intervention. This precluded assessment of balance at baseline and adjustment for baseline values in the analysis, e.g. through difference in difference. While our comparison groups were well-balanced overall, and we adjusted for unbalanced covariates, we cannot account for unobserved confounding. Collecting outcome data four years after the intervention without a baseline also precluded identification of any secular trends. There has been migration into and out of the setting, as well as autonomous upgrading amongst some control compounds. We excluded people who had lived on the compound for less than four years. However, risk of bias remains, for example if people with certain characteristics were more likely to have moved away after new toilets were constructed.

Since SanQoL data are necessarily subjective, the risk of reporting bias is difficult to avoid. We assume that any measurement error was not correlated with toilet type. However, there could have been social desirability bias, for example those having received the intervention may have wanted to appear grateful. Those who didn't receive it may have wanted to appear badly off, in the hope of receiving a future intervention. Therefore, the strongly positive effect on the primary outcome could be biased upwards. With a sample size of 424, we were underpowered for sub-group analysis. This limits the conclusions to be drawn from insufficient evidence for gender and age interactions.

While the QoL concepts addressed in SanQoL questions (e.g. safety, health risk) have broader relevance for QoL, the questions only focus on their manifestation in relation to sanitation (Table 1). This limits the conclusions we are able to draw in relation to QoL more broadly, except for the good evidence for a positive association with mental well-being as measured by the WHO-5. All our respondents had lived on the same compound before the intervention so had experience of low-quality pit latrines. However, their previous experiences are likely to have influenced how they scaled their responses to SanQoL questions. For example, the types of toilets used earlier in life, and whether they had always lived on the compound or migrated from elsewhere, might have influenced baseline considerations of what is disgusting or what level of privacy is acceptable.

To our knowledge this study represents the first use of a VAS in investigating people's self-perceived level of sanitation, so further investigation is required of its properties, including qualitative work exploring how people approach the exercise. Before applying the sanitation VAS in economic evaluation, due consideration will be need to be given to the extent to which advantages and disadvantages of VAS as applied in health economic evaluation are relevant (Parkin and Devlin, 2006; Torrance et al., 2001).

Conclusion

We estimated the impact of a shared sanitation intervention on quality of life (QoL). Our first result is that the intervention was associated with a 1.6 standard deviation increase in sanitation-related quality of life ($p<0.001$). This difference was seen across all five SanQoL attributes (disgust, health, safety, shame, and privacy) but with particularly large effect sizes for privacy and safety. A similar effect size was identified when sanitation status was measured by visual analogue scale (VAS) scores instead of psychometric items. This is the first exploration of the QoL effects of a specific sanitation intervention.

Our second result is that using a pour-flush toilet as opposed to a pit latrine was associated with a 0.2 standard deviation improvement in mental wellbeing, as measured by the WHO-5 index (p=0.04). The smaller effect size is expected, since toilet type is only a minor determinant of mental wellbeing but a major determinant of SanQoL. Previous studies have identified associations between sanitation services and mental wellbeing in cross-section (Caruso et al., 2018; Gruebner et al., 2012). However, ours is the first we know of to evaluate a specific intervention.

These two findings demonstrate that better toilets can improve people's lives beyond reducing infectious disease, at a time when several randomised trials have called the short-term health effects of incremental sanitation improvements into question. It was unsurprising that people using better-quality toilets would experience better privacy, less disgust, and so on. However, in terms of policy relevance, our first contribution is in quantifying the degree of this difference. Such quantifications can inform decisions based on comparative effectiveness. In particular, economic evaluations employing cost-effectiveness and cost-benefit analysis can value these differences and inform efficient allocation of resources. QoL benefits have never been incorporated into economic evaluations of sanitation programmes, and excluding them risks undervaluing sanitation interventions.

An additional contribution is in demonstrating that shared sanitation can be high-quality and improve QoL. Sanitation facilities shared by more than one household are still considered "limited" in global monitoring (UNICEF & WHO, 2019). In urban LMIC settings, open defecation is usually low, and the principal problem is how to move people from low-quality to high-quality sanitation. Our results are therefore relevant for decisions about sanitation in high-density settings with tenure and space constraints, where shared toilets may be the only realistic medium-term solution.

Future intervention trials might include wellbeing and QoL outcomes, to enable assessment of the comparative effectiveness of interventions on this basis, alongside other outcomes such as disease. Future studies of QoL effects of sanitation interventions would ideally be adequately powered for exploring effects by gender and age. Policy-makers may be willing to make decisions on the basis of QoL outcomes valued by users. These benefits should also be more achievable and measurable in the short-term, even under uncertainty about health benefits.

Funding. 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 interest. None declared.

Acknowledgements. We greatly appreciated the cooperation of survey respondents. We also appreciated the efforts of the fieldworkers employed by WE Consult: 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 benefitted from the comments of Roxanne Kovacs and seminar participants at the Centre for Health Economics in London at LSHTM, as well as from Britta Augsburg and Antonella Bancalari.

Chapter 8: Introducing the sanitation-adjusted person year (SAPY)

8.1. Introduction

Chapters 5-7 of this thesis have developed a method for measuring sanitation-related quality of life (SanQoL). This could be useful in measuring the impact of interventions on the attributes of sanitation which toilets users think are most important. However, to include these benefits in economic evaluation, they must not only be measured but also valued.

In this chapter, I propose an extra-welfarist measure of the value of sanitation. Named the sanitation-adjusted person year (SAPY), it represents a year of sanitation service experience at full sanitation-related capability. In doing so, I build on theory underlying the quality-adjusted life year (QALY), though a fundamental difference is that the 'length' dimension is years experiencing a particular service rather than life years. This chapter describes the rationale for the SAPY, its theoretical properties, and its limitations. Chapter 9 then applies the SAPY in a cost-effectiveness analysis, using the SanQoL effect estimates from Chapter 7. Further discussion of the SAPY is undertaken in chapter 10 once those empirical results have been presented. Since chapter 9 uses the SAPY in a CEA study, the focus in this chapter is on the SAPY as potentially used in that way. However, given the preference in the sanitation sector for welfarist CBA, it may be important to develop methods for estimating the monetary value of a SAPY, which I discuss in Chapter 10. In doing so, it will be necessary to overcome challenges in reconciling the capability approach with welfarism.

8.2. Distinguishing sanitation services, sanitation states and SanQoL index values

Before explaining the SAPY, it is important to distinguish three things: someone's level of sanitation service, their sanitation state, and the SanQoL index value of that state. Furthermore, it is important to understand the source of valuation (e.g. the person experiencing the state or the general public).

Consider a middle-aged man and his adolescent daughter, who both use a good-quality pour-flush toilet in their yard, shared with their neighbours. They both have a "limited" level of service in SDG terms. However, their sanitation states could be markedly different, depending on conversion factors (see conceptual model in Chapter 5). Assume that the door does not shut or lock properly. The father may not mind this, and might be in the best possible sanitation state: 11111. Conversely, the daughter may mind, and may therefore report worse levels for SanQoL privacy and safety items, say 11312 (Table 1). This illustrates the difference between the two individuals' levels of service, which are the same, and their sanitation states (or sanitation capability states), which are different.

Table 1: Descriptive system for the SanQoL measure

	Attribute	Questionnaire item	Responses*
1	**Disgust**	Can you use the toilet without feeling disgusted?	Always (1) Sometimes (2) Rarely (3) Never (4)
2	**Health**	Can you use the toilet without worrying that it spreads diseases?	
3	**Privacy**	Can you use the toilet in private, without being seen?	
4	**Shame**	Can you use the toilet without feeling ashamed for any reason?	
5	**Safety**	Are you able to feel safe while using the toilet?	

*respondents can choose "prefer not to answer" for any item.

When it comes to valuation, assume that the rank sum approach set out in Chapter 6 is being used. This will be improved upon in the future, and the system of item scores and weights replaced with a value set approach similar to EQ-5D and ICECAP-A. Recall that 11111 is the best state and 44444 the worst, and for now a level of 1 receives an item score of 3, and a level of 4 receives an item score of 0, and so on, as shown in Chapter 6. Using the weights in Chapter 6, the SanQoL index value of the father's state (11111) is 1.00 and of the daughter's state (11312) is 0.82 (Table 2).[5] It is logical that the value of 11312 would be lower than 11111, since several of the attributes are at lower levels.

[5] State 11111 calculation: 1.00 = (3*0.22)+(3*0.29)+(3*0.20)+(3*0.13)+(3*0.17)
State 11312 calculation: 0.82 = (3*0.22)+(3*0.29)+(1*0.20)+(3*0.13)+(2*0.17)

Table 2: Scoring for three sanitation states

	Weights and item scores					Index value
	Disgust weight	Health weight	Privacy weight	Shame weight	Safety weight	
	0.22	0.29	0.20	0.13	0.17	
Sanitation state	Disgust item scores	Health item scores	Privacy item scores	Shame item scores	Safety item scores	
11111	3	3	3	3	3	**1.00**
11312	3	3	1	3	2	**0.82**
31112	1	3	3	3	2	**0.81**

Assume a third person is in state 31112. It is less obvious whether they would have a higher or lower index value than the daughter (11312). Both of these states have three attributes at “always”, one at “sometimes” and one at “rarely”, but for a different set of attributes. As shown in Table 2 the index value of 31112 is lower (0.81) than 11312 (0.82), driven by the higher weight for the disgust attribute which is now at “rarely”. This illustrates the difference between sanitation states and index values.

The question of whose values shape the mapping of states onto index values is important – they could be based on the views of the general public or the people experiencing the state (Brazier et al., 2018). The weights in Chapter 6 were derived from a survey of the study population, but for now let us assume they represent “general public” weights. Given the structure of the SanQoL dataset, however, it would also have been possible to use the individual’s own ranking of the SanQoL attributes to calculate their private set of weights and then calculate their individually-weighted index value. I did not do this because, to inform the use of public funds, the general public is usually the preferred source of valuation (Brazier et al., 2016). Trade-off studies used for valuation report standard errors of estimates, allowing economic evaluations to take account of uncertainty around valuation in probabilistic sensitivity analysis. It would also be possible to use the sanitation visual analogue scale (VAS) scores (Chapter 6) to weight SAPYs. Doing so would be make the valuation of different dimensions of sanitation implicit and individual, rather than being derived from psychometric items. It would also be another means of experience-based valuation.

8.3. Rationale for a sanitation-adjusted person year (SAPY)

QALYs and DALYs are types of health-adjusted life years (Gold et al., 2002). There is no reason why life years (if anchored at death) or other measures of time cannot be weighted with other outcomes. The idea of extra-welfarist measures for sectors other than health has been raised before (Broome, 2004). Weatherly et al. (2009) propose that "crime QALYs" or "education QALYs" could be used in the economic evaluation of public health interventions, while Brouwer et al. (2008, p. 335) imagine their use in "education, housing, criminal justice" .

Indeed, education economists focused on LMICs have already developed the concept of "learning-adjusted years of schooling" (LAYS), which weight years of schooling by measured learning outcomes (Filmer et al., 2020). So far, these have only been applied as "macro-LAYS" at the national level, but some have proposed the micro-LAYS which could be used to evaluate interventions (Angrist et al., 2020).

Some have proposed "person-years of latrine access" (International Rescue Committee, 2016) or "water-person-years" (Koestler et al., 2010). However, these have been a simple multiplication of users and infrastructure useful life. This includes no weighting according to the quality of life of users, or even the quality of the service. Therefore, combining them with cost data would remain a cost analysis rather than a CEA.

The rationale for the SAPY is that it could be used in CEA or CBA. In CEA, it could be used as an extra-welfarist measure of value, as part of a comparison of incremental costs and incremental SAPYs. In welfarist CBA, the monetary value of a SAPY could be estimated using willingness to pay methods. Alternatively, the SAPY could be used as part of a cost-consequences analysis of multi-criteria decision analysis (Chapter 2). Any of these approaches would comprise including quality of life benefits of sanitation in economic evaluation for the first time.

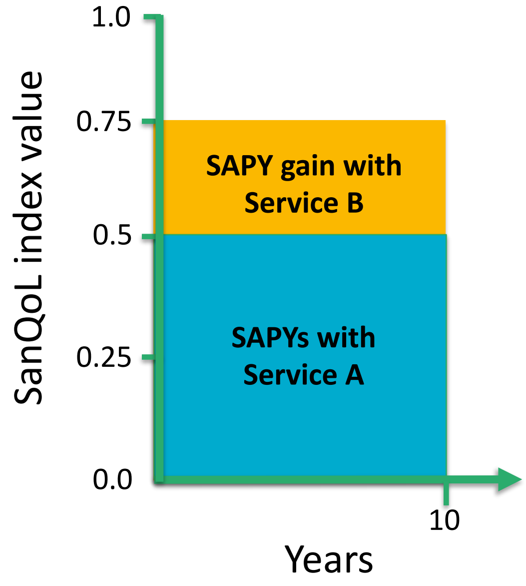
8.4. Properties of the proposed SAPY

The SAPY is a measure of the value of sanitation. One SAPY represents a year of sanitation service experience at full sanitation-related capability. SAPYs are calculated by weighting time experiencing sanitation services with an individual's level of SanQoL during that time. The approach to anchoring builds on the ICECAP measures' framing of anchor points from a philosophical perspective, whereby the best and worst states in the descriptive system are anchored at either end of the 0-1 scale (Coast et al., 2008a). The 'length' dimension in the SAPY comprises person-years experiencing a particular service, rather than life years. Another way of thinking about SAPYs could be as "years of full [sanitation-related] capability equivalent" (Goranitis et al., 2017) but this would be harder to explain to intended users of the measure in the sanitation sector.

Since they are anchored and weighted in different ways, a SAPY is not equal in value to a QALY. Likewise, it is not equal to a "year of full capability equivalent" because the evaluation is restricted to sanitation-related capabilities, not overall capabilities. SAPYs have no conceptual link to length of life or death. People with no sanitation capability generate no SAPYs. This might be the case if they have the lowest possible level of SanQoL, as well as if they are in a coma or dead, but "dead" is not a state within the SanQoL descriptive system.

For example, 1 SAPY is worth one year using a service where the user's SanQoL = 1, but 1 SAPY is also worth two years where SanQoL = 0.5. SAPYs allow different levels of SanQoL for different periods of time and/or at different times to be traded off. For example, five person-years could be spent in two different ways: (i) two years at high SanQoL followed by three years at low SanQoL, (ii) five years at medium SanQoL. A comparison could be made between SAPYs associated with these two options and their relative costs. A toilet used by five people all at full sanitation capability would be associated with five SAPYs per year. Figure 1 provides a visual explanation, in which two hypothetical sanitation technologies are compared. Both have a useful life of 10 years, but Service A provides the user a SanQoL of 0.5 while Service B provides 0.75. Service A generates 0.5 x 10 = 5 SAPYs to each user while Service B generates 0.75 x 10 = 7.5 SAPYs, representing a 'SAPY gain' of 2.5 per user over a 10-year horizon. For the five-user population, the SAPY gain would be 12.5.

Figure 1: Visualising SAPY gain as a result of an intervention



Assume that a decision-maker is comparing Service A and Service B as possible interventions in a population currently practising open defecation. Assume also that Service B costs slightly more than Service A. If the decision-maker cannot perceive any qualitative difference between the two options, then they might choose Service A on a purely cost-based comparison. However, if they undertook a CEA comparing incremental SAPYs (after discounting) against the incremental discounted costs of Service A, the decision might switch, depending on their willingness to pay for SAPY gains. This scenario is explored in an empirical CEA in Chapter 9.

8.5. Discussion

To my knowledge, the SAPY is the first user-weighted measure of the value of sanitation. It could be used in extra-welfarist CEA, or in CBA after monetary valuation. Just as a QALY can be used to compare interventions for arthritis to those for Zika virus disease, a SAPY can be used to compare community-led total sanitation to public toilets to hardware subsidies in a given setting. SAPYs are a measure of the value of sanitation as a state of being. Everyone defecates and urinates, and within a household, individuals often use the same toilet. However, there can be intra-household variation in SanQoL due to heterogenous experience of the service. The focus on the user experience makes SAPY a measure of outcomes rather than infrastructure or service quality, just as with the "natural units" sometimes used in health CEA (Drummond et al., 2015). It incorporates the useful life of technologies, allowing time at a given level of SanQoL to be traded off with QoL gains. SAPYs could also be used to explore equity questions, for example using equity impact planes or exploring opportunity costs by wealth quintile, as has been done for QALYs (Cookson et al., 2017).

The evaluative space of SAPY-based CEA is “sanitation” as measured by capability-based scaling of SanQoL. This means pursuing and valuing sanitation as a state of being for its own sake, rather than because it brings utility, just as QALY-based CEA aims to maximise health rather than the utility derived from health (Brouwer et al., 2008). While inter-personal comparisons of utility are not permitted under welfarism (see Chapter 2), use of SAPYs in an extra-welfarist CEA permits inter-personal comparison of levels of sanitation. However, it would be important to consider in future research whether to treat sanitation as a maximand (as in QALY-based CEA) or instead to draw on the “sufficient capability” concept (Mitchell et al., 2015). This might involve comparing “years of sufficient sanitation capability equivalent” rather than SAPYs, which are akin to “years of full sanitation capability equivalent” (Goranitis et al., 2017). To do so would require further study of what people in the setting consider “sufficient” in the SanQoL descriptive system (Kinghorn, 2019). For now, then, the analysis in Chapter 9 implicitly aims to maximise sanitation-related capability.

Estimating the SAPY gains from sanitation interventions should be more straightforward than estimating DALYs averted, for two reasons. First, calculating SAPYs only requires data on SanQoL, the estimated useful life of the technology and the mean number of users. Effects on SanQoL should be less difficult and expensive to estimate than for effects on diarrhoea or other health outcomes, due to the simplicity of the questions and arguably diminished risk of confounding. The effect estimates required to calculate DALYs rarely exist at the level of specific sanitation interventions rather than broad levels of service, and are very expensive to estimate (Wolf et al., 2018). Second, regardless of the study design used for causal inference, the plausibility in attributing a quality of life effect to a sanitation intervention is higher than for an effect on diarrhoea or other health outcomes. This is because there likely to be far fewer confounding factors than the multiplicity of causal pathways that exist for diarrhoeal pathogens (Wagner and Lanoix, 1958). In other words, there are fewer other important factors in the causal chain between toilets and sanitation-related disgust than there are between toilets and diarrhoea. It is possible to imagine SanQoL data collection becoming routine in programme monitoring and evaluation in a way that is not true for health outcomes. This could enable relatively low-cost economic evaluations, though robust causal inference would still be required.

Sanitation interventions can prevent diarrhoeal morbidity and, by extension, mortality (Wolf et al., 2018). To incorporate both quality of life effects and disease risk reduction effects in an economic evaluation, using CBA would likely be appropriate. SAPY-based CEA may still play a role for some decisions, for example where there is low likelihood of health and time savings benefits, such as the scenario in Chapter 9. While health regularly appears in lists of important characteristics of overall well-being and quality of life (Nussbaum, 2011; Stiglitz et al., 2009), it is harder to make the case for sanitation appearing in such lists. Nonetheless, sanitation as a state of being is valued by individuals. Furthermore, in 2015 the UN General Assembly recognised the human right to sanitation as a distinct right (United Nations, 2016b). Governments in democracies likely invest in sanitation because they value health and its impact on productivity, but also because they know that citizens seeing improvements in the cleanliness of their neighbourhoods are more likely to support a government which achieves this.

Note that “no sanitation service” (i.e. open defecation) is unlikely to mean “no SAPYs”. This is because the mean SanQoL index value in a population practising OD is unlikely to be zero. In some settings, it is common to have a revealed preference for OD over a low-quality toilet (Coffey et al., 2014), implying higher SanQoL from OD. Even in the absence of a preference for OD over toilets, some people may not find the practice particularly disgusting, unsafe, and so on. Any level of service, including OD, will not be associated with a uniform level of SanQoL, due to conversion factors. Practising OD in a densely-populated urban area may be associated with lower SanQoL than in a rural area.

To incorporate SAPYs, CBA studies could elicit willingness to pay for them. This could be done using some of the methods applied in estimating the monetary value of a QALY (Ryen and Svensson, 2015), for example contingent valuation (Bobinac et al., 2010; Gyrd-Hansen, 2003). The ensuing monetary values could be applied in CBA without double-counting, since SanQoL measures capture benefits not yet included (Hutton and Chase, 2016). The SanQoL health question focuses on the value of avoiding worry about disease, rather than the value of actual disease risk reductions.

There are some limitations to the SAPY and the SanQoL measure underlying it. First, people's subjective valuations of their present states of being are affected by their previous experience, a phenomenon called adaptation (Nussbaum, 2001; Teschl and Comim, 2005). Most people are likely to have experienced something close to "full health" (HRQoL = 1) for some part of their life. However, this is less likely to be the case for SanQoL, perhaps making adaptation a more important consideration. For example, a renter who has moved from dwelling with a high-quality toilet to one with a low-quality one may rate their SanQoL lower than someone who has only ever used a low-quality toilet.

Second, the valuation method used in Chapter 6, based on the rank sum method and attribute scoring, does not strictly speaking meet the Brazier et al. (2016) conditions for use of measures to weight outcomes in economic evaluation. This is because, while the source of valuation is from the users, it does not come from a trade-off method. However, this work is preliminary and illustrative for a first exposition of the SAPY approach. A trade-off study using methods such as discrete choice experiment or best-worst scaling could be undertaken in the future. Using such a value set might change the empirical findings reported in other chapters.

Third, as with measures such as EQ-5D and ICECAP-A, SanQoL necessarily has a small number of items, in order to make future stated preference trade-off studies viable (Hensher et al., 2015). This means that broad concepts are measured, e.g. "disgust" rather than more specific ones such as seeing insects, smelling faeces, etc. This risks the perception that some concepts are "excluded", though a more accurate characterisation would be "subsumed". To that extent, the fact that SanQoL focuses on a small number of high-level concepts is a limitation, just as it is for all health economic measures used in economic evaluation. However, the attributes identified in the qualitative research (Chapter 5) are the core concepts of SanQoL in this setting, which is best practice in measure development (Coast et al., 2017). Developers of short-form measures typically aim for high correlation with long-form measures of the same construct (Brazier et al., 2016). Such measures did not exist for me to compare against. However, if longer-form measures of SanQoL are developed in the future, this could be explored.

Fourth, SAPYs only value the user-experienced benefits of the part of sanitation related to toilet use. They do not capture welfare gains from cleaner neighbourhood environments, the quality of life of sanitation workers, and so on. Other methods would be needed to measure and value these benefits. As such, SAPYs are unlikely to capture effects of improving faecal sludge management or wastewater treatment services in ways which do not directly benefit individuals within the five SanQoL attributes.

Fifth, the SAPY is quite sensitive to assumptions about useful life of assets, which could be problematic if options being compared have very different useful lives. Since useful life is hard to explore empirically without revisiting assets 10-20 years after an intervention, it is an important assumption to be tested in sensitivity analysis.

8.6. Conclusion

This chapter proposed that the sanitation-adjusted person year (SAPY) could be used in CBA and CEA as a measure of the value of sanitation to toilet users. When used in extra-welfarist CEA, SAPYs offers a more meaningful way to compare the economic performance of interventions than cost analysis alone, but exclude other benefits beyond quality of life. When used in CBA, SAPYs allow the valuation of benefits that have hitherto been excluded. Future research priorities are further investigation of the validity and reliability of SanQoL in different settings, as well as valuation studies using trade-off elicitation methods to replace the rank-sum method used in Chapter 6.

Chapter 9: Research Paper 5 - Cost-effectiveness analysis

Chapter introduction

In Chapter 3 I demonstrated that the quality of life benefits of sanitation have never yet been included in an economic evaluation of a sanitation intervention. In Chapters 5-6, I developed a SanQoL measure to capture these benefits and value them on a 0-1 scale. In Chapter 7, I estimated the effect of an actual sanitation intervention in Maputo, Mozambique on SanQoL index values. In Chapter 8, I proposed the sanitation-adjusted person year (SAPY) as an extra-welfarist measure of the value of sanitation.

In this chapter (Research Paper 5), I apply the effect estimate from Chapter 7 in a cost-effectiveness study of the intervention evaluated by the MapSan trial, using the SAPY as the measure of the value of sanitation. This achieves objective 4 of this thesis: to estimate the cost-effectiveness of a shared sanitation intervention by valuing quality of life gains. I am proposing to submit the paper to *Health Economics* (published by Wiley). It follows the standard structure for biomedical journals: introduction, methods, results and discussion. Supplementary Materials are in Appendix E at the end of this thesis.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1701671	**Title**	Mr
First Name(s)	Ian		
Surname/Family Name	Ross		
Thesis Title	Measuring and valuing quality of life in the economic evaluation of sanitation interventions		
Primary Supervisor	Giulia Greco		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Health Economics
Please list the paper's authors in the intended authorship order:	Ian Ross, Catherine Pitt, Giulia Greco, Zaida Adriano, Rassul Nala, Joe Brown, Oliver Cumming
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceptualised the study, developed the methods, and coordinated inputs from the supervisory team. I trained fieldwork teams and participated in piloting and data collection. I conducted the analysis and write-up, revising drafts on the basis of co-author feedback.

SECTION E

Student Signature	[redacted]
Date	24/10/20

Supervisor Signature	[redacted]
Date	24/10/20

Using sanitation-adjusted person years to assess quality of life effects of sanitation interventions: a novel cost-effectiveness analysis

Authors:

Ian Ross[1], Catherine Pitt[1] , Giulia Greco[1] , Zaida Adriano[2], Rassul Nala[3], Joe Brown[4], Oliver Cumming[1]

Affiliations:

[1]London School of Hygiene and Tropical Medicine, UK;

[2]WE Consult, Mozambique;

[3]Instituto Nacional de Saúde, Mozambique;

[4]Georgia University of Technology, USA;

Corresponding author:

Ian Ross

LSHTM, 15-17 Tavistock Place, London, WC1H 9SH

Email: ian.ross@lshtm.ac.uk

Phone: 020 7636 8636

Abstract

Several recent randomised trials found mixed evidence for the effect of basic sanitation interventions on childhood diarrhoea. Sanitation interventions can improve quality of life outcomes not captured in health value frameworks, such as privacy, safety and dignity, which are particularly valued by users. We demonstrate a novel approach to extra-welfarist cost-effectiveness analysis of sanitation interventions using sanitation-adjusted person years (SAPYs). The Maputo Sanitation Trial in urban Mozambique concluded that there was no evidence that a shared sanitation intervention affected childhood diarrhoeal disease. However, people's quality of life and mental wellbeing improved. Characteristics of the provided assets beyond the user interface were likely important for quality of life, such as the solid walls and lockable door. We estimated the incremental costs and outcomes of shared pour-flush toilets (STs) and community sanitation blocks (CSBs), relative to existing use of shared pit latrines (PLs), over a 15-year horizon. For outcomes, we used a novel measure of sanitation-related quality of life to value SAPYs. Relative to PLs, STs cost an additional $52 (95% CI: 40-66) per SAPY gained. CSBs were dominated. This demonstrates the usefulness of the SAPY in assessing quality of life gains from sanitation interventions. Both SAPYs and disease impacts could be valued in cost-benefit analysis, making monetary valuation of the SAPY a future research priority. Such extra-welfarist measures may be useful in other areas of public health.

Introduction

Sanitation is the separation of human excreta from human contact (WHO, 2018). Globally, 230 million people living in urban areas are without basic sanitation (UNICEF & WHO, 2019). Containing and treating excreta creates positive health and environmental externalities for the rest of society (Andrés et al., 2017). Systematic review evidence suggests that sanitation interventions can prevent diarrhoeal disease, with larger effects for higher levels of service (Wolf et al., 2018). However, several recent randomised trials found only mixed effects of basic sanitation interventions on childhood diarrhoea, and no effect on stunting (Humphrey et al., 2019; Luby et al., 2018; Null et al., 2018). The consensus amongst epidemiologists is that basic services alone are unlikely to have a large short-term impact on health outcomes, with more advanced services and complementary strategies required (Cumming et al., 2019).

A systematic review of the relationship between sanitation and mental well-being identified privacy and safety as root dimensions, alongside the related concepts of shame, anxiety, fear, assault, dignity and embarrassment (Sclar et al., 2018). Beyond health, cleanliness and convenience are also commonly reported as important by toilet users (Novotný et al., 2018). All of these are attributes of sanitation which users "have reason to value" (Sen, 1999) and therefore contribute to quality of life (QoL) under a capability approach (Sen, 1980).

Economic evaluations aim to inform investment decisions by comparing costs and outcomes of interventions, using methods such as cost-effectiveness analysis (CEA) and cost-benefit analysis (CBA) (Drummond et al., 2015). CEA within a health value framework provides only a limited picture of sanitation interventions' performance (Briscoe, 1984). While the health-related QoL measures underlying quality-adjusted life years (QALYs) usually include an item related to anxiety or depression (Brazier et al., 2016), this is unlikely to capture QoL effects of sanitation interventions in a meaningful way. Disability adjusted life years (DALYs) also do not capture the relevant outcomes. CBA studies, meanwhile, have thus far excluded QoL benefits as "intangible" (Chapter 4). Such exclusion risks misallocating the $30 billion invested annually in sanitation in low- and middle-income countries (LMICs) (WHO, 2017).

The first step towards including QoL effects of sanitation interventions in economic evaluation is to measure them. A recent study used qualitative research to develop a conceptual model of sanitation-related QoL (Chapter 5). A subsequent study developed a user-reported “SanQoL” measure and assessed its validity and reliability in an urban Mozambican setting (Chapter 6). However, QoL outcomes have not yet been valued in an economic evaluation.

We aim to demonstrate a novel approach to cost-effectiveness analysis of sanitation interventions by valuing quality of life effects using sanitation-adjusted person years (SAPYs). We worked alongside the Maputo Sanitation (MapSan) Trial in Mozambique (clinicaltrials.gov, NCT02362932), which concluded that a shared urban sanitation intervention had no impact on under-5 health outcomes (Knee et al., 2020). Depending on incremental costs, a CEA using DALYs would be likely to conclude that the intervention was not cost-effective. We explore the use of a QoL value framework, which could inform future sanitation investment decisions made by the Maputo municipality and beyond.

Methods

Intervention and setting

Of Maputo’s 1.1 million population (INE, 2019), 70% live in informal settlements (UN-HABITAT, 2010), and 89% use non-sewered sanitation facilities (Hawkins and Muximpua, 2015). The study setting is low-income neighbourhoods (*bairros)* in Maputo, Mozambique, comprising an approximately 10km^2 area of the Nhlamankulu district. In this area, it is common for households to live in informal multi-household ‘compounds’, sharing the same toilet. Such toilets are often self-built pit latrines (PLs), also used for bathing. These vary widely in their quality, but many comprise an informally-walled area of a compound rather than a building, providing little privacy (Capone et al., 2020). Such toilets have no roof, a makeshift door, soil or wooden floors, and no pit lining. Due to the ‘direct drop to pit’ interface without a water seal, they often smell bad (Shiras et al., 2018). When such pits fill, users sometimes cover them and dig a new one, or empty them if there is no space for this (Capone et al., 2020).

During 2015-16, a shared toilet installation project was implemented by Water and Sanitation for the Urban Poor (WSUP), a non-government organisation (NGO), as part of a broader programme funded by the World Bank. The intervention provided a pour-flush toilet discharging to a septic tank with soakaway. There were two design types, depending on user numbers. The first type consists of a single-stall shared toilet (ST) to be used by around 15 people, with users asked to contribute 15% of the construction cost, implying 85% subsidy. The second type is a Community Sanitation Block (CSB) with two stalls, to be used by at least 21 people, with users contributing 10%, implying 90% subsidy. By contrast to PLs, the ST and CSB options are built of cement blocks, have a metal roof, concrete floor, and a door which locks from the inside. The 'pour-flush' interface prevents smells. CSBs also included a handwashing basin, rooftop water tank and two laundry stations

Photographs of all three types are provided in supplementary material A, alongside more information on intervention activities and stakeholders. In summary, the implementing NGO contracted eight community-based organisations to engage potential users, select sites and collect financial contributions. Nine small local construction firms were contracted to prepare sites and build toilets. Oversight was provided by the World Bank and the *Conselho Municipal de Maputo* (CMM), the municipality. The intervention was delivered to 450 compounds. CSBs were installed in 50 compounds and STs in 400.

Study design

We sought to inform a decision between three alternatives: construct ST, construct CSB, or do nothing, resulting in households continuing to construct PLs. We used a decision analytic model based on a simple three-pathway decision tree without probabilities (diagram in supplementary material B). We estimated one set of costs and outcomes per pathway, modelled for a cohort of 7,200 people per alternative, which was the approximate number of people benefitting from the actual intervention as a whole.

We took a societal perspective with an "analytic horizon" (O'Mahony et al., 2015) of 15 years, which is the estimated average useful life of STs and CSBs. Since mortality and health outcomes were not considered, no meaningful differences in costs and effects between alternatives were anticipated after this time. PLs have has an estimated average useful life of 5 years, so we assumed a new PL to be built three times during the 15-year time horizon.

We collected cost data in a household survey in 2018, and outcome data in a second survey in 2019. Characteristics of respondents the 2019 survey were presented in Chapter 7. They were adults with a mixed age range, and about half were women. A quarter rented their dwelling, and half owned a fridge. Two-thirds of dwellings had concrete external walls. Characteristics of 2018 survey respondents are reported elsewhere (Capone et al., 2020), but are broadly similar, being from the same population.

Ethics

All survey participants were aged over 18 and provided written informed consent prior to data collection. Ethical approval was granted by the Research Ethics Committee of the London School of Hygiene and Tropical Medicine (Ref: 14609) and the *Comité Nacional de Bioética para a Saúde* (IRB00002657) at the Ministry of Health in Mozambique. Our methods and results are reported in accordance with the CHEERS consolidated health economic evaluation reporting standards (Husereau et al., 2013) – see supplementary material E – and the iDSI reference case for economic evaluations (Wilkinson et al., 2016).

Effects

The outcome for which effects were compared was the sanitation-adjusted person year (SAPY), an extra-welfarist measure of the value of sanitation. The SAPY is a novel measure which functions similarly to the QALY (Drummond et al., 2015). A year of a person's sanitation experience is weighted with sanitation-related quality of life (SanQoL) experienced during that year (Chapter 8). SanQoL builds on the health-related quality of life (HRQoL) concept integral to the QALY (Karimi and Brazier, 2016), but is based on the capability approach (Sen, 1985). SanQoL index values are measured on a 0-1 scale, where zero (one) represents no (full) sanitation-related capability (Coast et al., 2008a). One SAPY is the value of a year of sanitation experience at SanQoL = 1. However, SanQoL = 0 is not equivalent to death, but to "no sanitation-related capability", and a SAPY is not equal in value to a QALY. A visualised numerical example is provided in supplementary material B.

To calculate SanQoL index values, we used a five-item descriptive system, which was developed based on qualitative research in this setting (Chapter 5). Its capability-oriented questions cover five attributes: disgust, health, privacy, safety and shame (Table 1). Investigations of the measure's validity and reliability are reported elsewhere (Chapter 6).

Table 1: Descriptive system for the SanQoL measure

	Attribute	Weight	Questionnaire item	Responses*
1	**Disgust**	0.22	Can you use the toilet without feeling disgusted?	Always (1) Sometimes (2) Rarely (3) Never (4)
2	**Health**	0.29	Can you use the toilet without worrying that it spreads diseases?	
3	**Privacy**	0.20	Can you use the toilet in private, without being seen?	
4	**Shame**	0.13	Can you use the toilet without feeling ashamed for any reason?	
5	**Safety**	0.17	Are you able to feel safe while using the toilet?	

*respondents can choose "prefer not to answer" for any item.

SanQoL data were collected in the 2019 survey from toilet users (n=424) living on intervention and control groups compounds of the MapSan trial. There were no missing SanQoL data. Respondents also ranked the five attributes using a visual analogue scale. SanQoL index values on a 0-1 scale were calculated by weighting psychometric responses using the rank sum method (Stillwell et al., 1981), fully discussed elsewhere (Chapter 6). The results of mixed effects regression models estimating the effect of access to STs and CSBs on SanQoL, adjusting for covariates, are reported elsewhere (Chapter 7).

We calculated SAPYs for each of the three alternatives over the 15-year horizon. We assumed that construction occurred at the end of year 0 and SAPYs were accrued from year 1 onwards. The same level of SanQoL is assumed to be sustained throughout the time horizon. This is because costs of maintaining and cleaning toilets, and emptying pits and septic tanks, are included on the cost side, so there is no reason to suggest toilet quality (and therefore QoL) would deteriorate. We relaxed this assumption in sensitivity analysis. All alternatives were associated with the same number of person-years over the time horizon, because mortality, population growth, and migration were not considered. All alternatives are, however, associated with different numbers of SAPYs due to varying SanQoL index values. Survey data suggested that all compound members used the same toilet (Chapter 7), so all individuals in the model accrue SAPYs associated with the toilet type in the compound. Only adults over 18 were sampled, so we assume that the same index values apply to under-18s.

Costs

We estimated the incremental economic costs of STs and CSBs as compared to PLs, from a societal perspective. We included costs along the sanitation service chain: containment, emptying, transport, and treatment. We took a 'real world' costing approach (Vassall et al., 2017), with capital costs incurred in 2015 (year 0) and annualised over their useful lives during 2016-2030 (Drummond et al., 2015).

Capital costs of CSBs and PLs were estimated from the implementing NGO's financial records in a top-down retrospective approach, using allocation rules based on interviews with programme staff. Programme costs borne by other stakeholders (e.g. World Bank and CMM) were estimated based on quantities and prices of resources used in a bottom-up approach. The time of World Bank and CMM staff was valued at the individuals' gross salaries including employer on-costs (supplementary material C). Capital costs of PLs, and recurrent costs for all options, were estimated bottom-up from the 2018 survey (n=779) – sampling is described elsewhere (Capone et al., 2020).

We interviewed households on MapSan intervention and control compounds about expenditure on emptying services, time spent cleaning toilets, and itemised expenditure on maintenance and cleaning products. From these data, we applied sample means per technology, including zeroes when respondents had not incurred a given cost, except in the case of emptying costs for STs and CSBs (Supplementary Material C). We assumed that formal service providers serving STs/CSBs would take faecal sludge to a treatment plant (Bauerl et al., 2016), and we estimated sludge volumes using benchmark accumulation rates (Strande et al., 2018). The opportunity cost of household time in toilet cleaning and meeting participation was assumed to be time spent in unwaged domestic labour (rather than waged employment). This time was given the same value for all individuals, specifically 50% (Whittington and Cook, 2018) of the minimum wage in Mozambique in 2015 (supplementary material C). To estimate sludge treatment costs, borne by the Maputo municipality (CMM), we applied the recurrent cost per m^3 for faecal sludge treatment plants in the 2015 Maputo sanitation master plan (AIAS, 2015). We assumed PLs incurred no treatment cost, since users universally covered full pits or manually emptied them with on-site burial or illegal dumping (Capone et al., 2020).

We calculated total economic costs as the sum of annualised capital costs and recurrent costs over the period, adjusted to constant 2015 US dollars ($). We modelled in $ because NGO financial data were in $ and comprised 70% of costs, but we also report results in MZN. All cost data collected in MZN were salaries and non-tradeable services (Turner et al., 2019a). Therefore, we first adjusted them to 2015 values using World Bank (2018) GDP deflator data for Mozambique, and then converted to constant 2015 $ based on the average 2015 exchange rate (Investing.com, 2019).

Cost-effectiveness

We discounted costs and SAPYs at 3% (Wilkinson et al., 2016), considering alternatives in sensitivity analysis. We conducted household survey analyses in Stata 16 (StataCorp, 2019) and other analyses in Excel (Microsoft, 2019). We made three pair-wise comparisons: ST to PL, CSB to PL, ST to CSB. For each, we calculated incremental cost-effectiveness ratios (ICERs) by dividing discounted incremental costs by discounted incremental outcomes over the time horizon (Supplementary Material B). We then identified the cost-effectiveness frontier, namely the set of points on the cost-effectiveness plane where the given alternative is the most cost-effective choice.

We undertook probabilistic sensitivity analysis (PSA) to characterise parameter uncertainty. We modelled 1,000 iterations of a Monte Carlo simulation, with input parameters varying simultaneously according to appropriate probability distributions (Table 2). We estimated mean ICERs by dividing mean incremental costs by mean incremental effects of the 1,000 iterations, and a 95% uncertainty interval based on iteration percentiles. We explored decision uncertainty using cost-effectiveness acceptability curves (CEACs), which show the probability that an option is cost-effective under different thresholds for the decision-maker's willingness to pay (WTP) for SAPY gains. We used deterministic sensitivity analysis (DSA) to explore the impact on ICERs of uncertainty and heterogeneity surrounding individual parameters and assumptions. We present a tornado plot indicating the magnitude of changes in the ICER for ST compared to PL when parameters are at high and low plausible values in one-way and two-way DSA (Table 2). We also undertook DSA for the other two comparisons (ST-CSB and CSB-PL).

Table 2: Model input parameters and DSA results per parameter

<table>
<tr><th></th><th colspan="5">Parameter values</th><th colspan="2">DSA results (ST vs. PL)</th></tr>
<tr><th>Parameter</th><th>Base case</th><th>Low</th><th>High</th><th>Distribution</th><th>Source & justification for low/high values</th><th>Low</th><th>High</th></tr>
<tr><td>Outcomes</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>SanQoL - PL</td><td>0.49</td><td>0.46</td><td>0.53</td><td>beta</td><td rowspan="3">PSA: 95% CI based on s.e. of regression coefficient, with beta distribution since SanQoL is bounded at 0 and 1. DSA: two-way with ST lower bound and PL upper bound and vice versa</td><td rowspan="2">42</td><td rowspan="2">68</td></tr>
<tr><td>SanQoL - ST</td><td>0.84</td><td>0.79</td><td>0.89</td><td>beta</td></tr>
<tr><td>SanQoL - CSB</td><td>0.78</td><td>0.70</td><td>0.86</td><td>beta</td><td>n/a</td><td>n/a</td></tr>
<tr><td>Costs (for 7,200 people)</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Annuitised capital cost - PL</td><td>2,586</td><td>2,068</td><td>3,103</td><td>gamma</td><td>PSA: 95% CI from household survey, with gamma due to strong right skew. DSA: ±20% two-way with ST</td><td rowspan="2">44</td><td rowspan="2">60</td></tr>
<tr><td>Annuitised capital cost - ST</td><td>101,111</td><td>80,889</td><td>121,333</td><td>uniform</td><td rowspan="2">PSA: uniform ±20% of NGO financial report values, in absence of knowledge of distribution. DSA: as above.</td></tr>
<tr><td>Annuitised capital cost - CSB</td><td>251,705</td><td>201,364</td><td>302,046</td><td>uniform</td><td>n/a</td><td>n/a</td></tr>
<tr><td>Household recurrent financial expenditure - PL</td><td>11,079</td><td>6,072</td><td>16,086</td><td>gamma</td><td rowspan="3">PSA: 95% CI from household survey, using gamma due to strong right skew. DSA: 95% CI in two-way with ST/PL recurrent cost</td><td rowspan="2">47</td><td rowspan="2">57</td></tr>
<tr><td>Household recurrent financial expenditure - ST</td><td>38,619</td><td>32,000</td><td>45,239</td><td>gamma</td></tr>
<tr><td>Household recurrent financial expenditure - CSB</td><td>64,586</td><td>46,797</td><td>82,376</td><td>gamma</td><td>n/a</td><td>n/a</td></tr>
<tr><td>Cleaning VOT (recurrent) - PL</td><td>7,614</td><td>3,807</td><td>11,422</td><td>n/a</td><td rowspan="3">DSA: 25% and 75% of wage as opportunity cost of time instead of 50% (Whittington & Cook, 2018), assessed one-way with both ST and PL varied simultaneously</td><td rowspan="2">51</td><td rowspan="2">52</td></tr>
<tr><td>Cleaning VOT (recurrent) - ST</td><td>9,828</td><td>4,914</td><td>14,741</td><td>n/a</td></tr>
<tr><td>Cleaning VOT (recurrent) - CSB</td><td>9,997</td><td>4,998</td><td>14,995</td><td>n/a</td><td>n/a</td><td>n/a</td></tr>
<tr><td>Sludge treatment cost (annual recurrent) - ST & CSB</td><td>1,808</td><td>1,175</td><td>2,440</td><td>n/a</td><td>DSA: Maputo sanitation master plan (AIAS, 2015) and Strande et al. (2018) - low and high are ±35% of base case with a uniform distribution, based on those two sources</td><td>52</td><td>52</td></tr>
</table>

Parameter	Base case	Low	High	Distribution	Source & justification for low/high values	Low	High
Other parameters							
Discount rate	3%	0.01%	10%	n/a	DSA: low case is close to zero, high case is Mozambican central bank base rate (Wilkinson et al., 2016)	44	75
Useful life - PL	5	7	3	n/a	Based on interviews with programme stakeholders, household survey and Hutton & Varughese (2016) assumptions	43	68
Useful life - ST	15	10	20	n/a			
Useful life - CSB	15	10	20	n/a		n/a	n/a
annual SanQoL decline	0%	n/a	2%	n/a	assumption, 2% for ST & CSB, 4% for PL (returns to full SanQoL each time a new PL is built)	n/a	70
2015 MZN / USD exchange rate	40.1	32.1	48.1	n/a	±20% on the base case	50	55

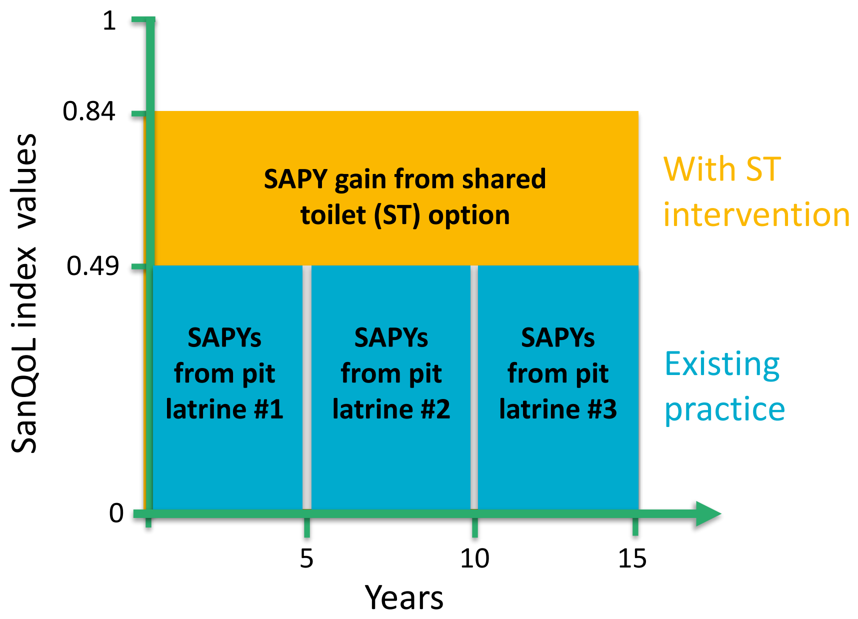
note. s.e. = standard error,

Results

Effects

SanQoL index values were the highest for ST and lowest for PL (Table 2). Delivering and sustaining sanitation services for 7,200 people over 15 years was associated with 42,392 SAPYs for PLs, 72,322 for STs and 67,343 for CSBs (Table 4). Figure 1 visualises SAPY gains in the base case for ST compared to PL.

Figure 1: Visualisation of SAPY gain over the 15-year horizon for ST compared to PL



Costs

Delivering and sustaining sanitation services for 7,200 people over 15 years costs $0.25 million for PL, $1.81 million for ST, and $3.92 million for CSB (Table 3). The economic costs of ST and CSB are therefore seven and 15 times higher than those of PL, respectively. For STs and CSBs, 'works' (materials and labour for construction) was the largest category of lifecycle economic costs, while for PLs 'cleaning materials' was the largest. Annualised economic cost per person was $2 for PLs, compared to $17 and $36 for STs and CSBs respectively (Table 4). Works costs per person of STs and CSBs were 40 and 130 times more than PLs, respectively, while recurrent costs per person of STs and CSBs were three and four times higher than PLs, respectively (Supplementary Material D).

Table 3: Categories of economic costs (in 2015 US$) of the three options

		Strategies					
		Pit latrine (PL)		Shared toilet (ST)		Community sanitation block (CSB)	
Cost categories		$	%	$	%	$	%
Capital	NGO staff (in Mozambique and UK)			383,595	21%	434,171	11%
	WB staff (in Mozambique)			38,649	2%	38,649	1%
	CMM staff (in Mozambique)			15,186	1%	15,186	0.4%
	NGO consultants (for site selection, design, works supervision)			210,364	12%	842,342	22%
	WB consultants (for monitoring)			11,220	1%	11,220	0.3%
	Works (ground preparation, toilet construction)	30,866	12%	461,280	26%	1,565,601	40%
	Time of *Chefe do Quarteirão*			105	0.01%	105	0.003%
	Household participation time			3,913	0.2%	3,913	0.1%
	NGO overheads (e.g. office, transport)			82,744	5%	93,654	2%
	Capital sub-total	**30,866**	**12%**	**1,207,056**	**67%**	**3,004,841**	**77%**
Recurrent	Cleaning expenditure	94,335	37%	410,270	23%	618,768	16%
	Cleaning value of time	90,901	36%	117,320	6%	119,344	3%
	Maintenance expenditure	8,120	3%	13,602	1%	124,517	3%
	Emptying expenditure	29,802	12%	37,163	2%	27,744	1%
	Treatment cost			21,580	1%	21,580	1%
	Recurrent sub-total	**223,158**	**88%**	**599,935**	**33%**	**911,952**	**23%**
Total		**254,024**	**100%**	**1,806,992**	**100%**	**3,916,793**	**100%**

A more disaggregated cost table is provided in Supplementary Material D. “Works” cannot be disaggregated into materials and labour.

Table 4: Outputs, outcomes, costs, and cost-effectiveness

		Pit latrine (PL)	Shared toilet (ST)	Community sanitation block (CSB)
Outputs	number of toilets	583	548	249
	mean users per toilet	12.4	13.1	29.0
	total people served	7,200	7,200	7,200
Outcomes*	total SAPYs	42,392	72,322	67,343
Total economic cost*	overall	254,024	1,806,992	3,916,793
	of which capital	30,866	1,207,056	3,004,841
	of which recurrent	223,158	599,935	911,952
Economic cost per person*	total cost per person	35	251	544
	annualised total cost per person	3	21	46
Cost-effectiveness	incremental cost effectiveness ratio (cost per SAPY gained, compared to PL)	.	52	dominated**

All costs in 2015 US dollars. * Summed over the 15-year time horizon and discounted. **Since CSB is dominated by ST, no ICER for CSB is on the cost-effectiveness frontier (Figure 2). In the specific circumstance where ST is infeasible on engineering grounds, for example, the ICER for CSB would be $147 (incremental to PL).

Cost-effectiveness

Base case incremental cost-effective ratios (ICERs) are reported in Table 3. In the probabilistic sensitivity analysis, the mean incremental cost per SAPY gained (ICER) for ST as compared to PL was US$ 52 (95% CI: 40 to 66). In the cost-effectiveness plane (Figure 2), datapoints represent the 1,000 simulations. Results for ST and CSB are presented as incremental to PL at the origin. No ST or CSB datapoints fall outside the north-east quadrant, implying that under all simulations these options deliver a SAPY gain over pit latrines, but at a higher cost. However, The CSB option was dominated by ST, in that its mean incremental costs were higher, and its mean incremental effects lower. The dotted line from the origin to mean values for ST therefore represents the cost-effectiveness frontier, the slope of which is the ICER. Figure 1 visualises SAPY gains in the base case. In the ST-CSB comparison, SAPYs were higher under ST in 89% of simulations, and costs lower under ST in 100% of simulations.

Figure 2: Cost-effectiveness plane showing parameter uncertainty around estimates of incremental costs and SAPYs as compared to PL

ST/CSB results are incremental to PL at the origin. The dotted line represents the cost-effectiveness frontier, the slope of which is the ICER.

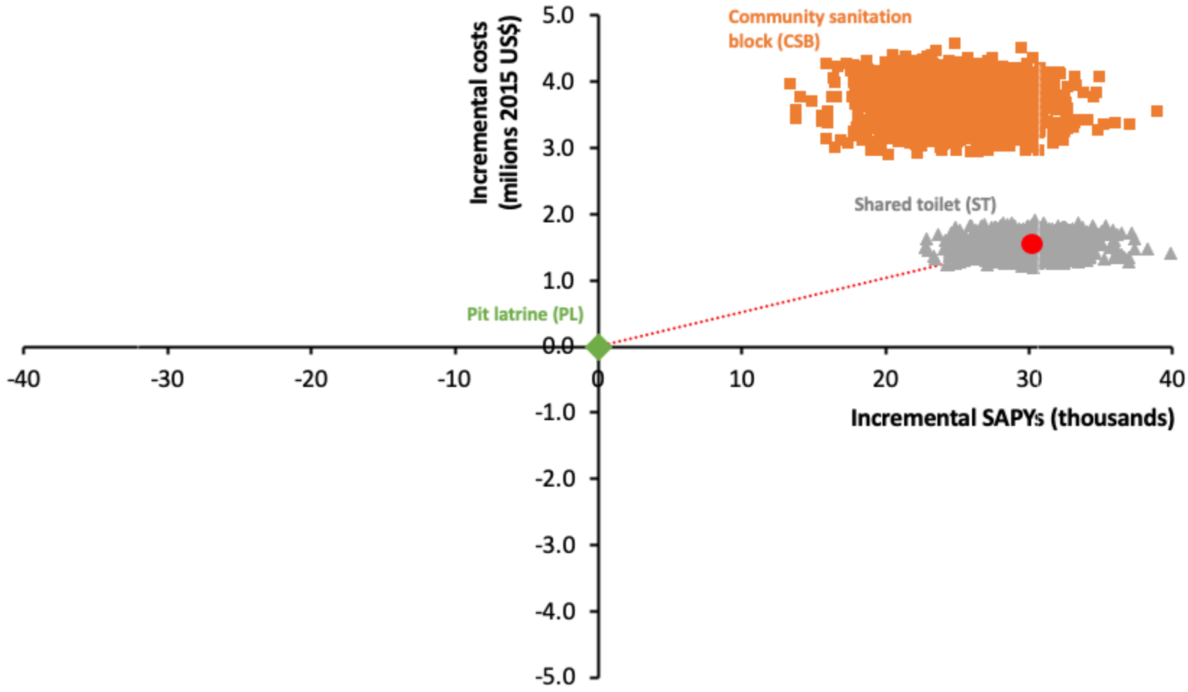


Figure 3 presents a cost-effectiveness acceptability curve (CEAC). The blue line illustrates the probability that the ST option is cost-effective (as compared to PL), under different thresholds for the decision-maker's willingness to pay (WTP) for SAPY gains. As the SAPY is a novel measure of the value of sanitation, there are no established WTP thresholds. Figure 3 illustrates that if the threshold were US$ 60 / SAPY, for example, there would be a 92% chance that ST was cost-effective. Since CSB is dominated, it has a 0% probability of being cost-effective.

Figure 3: Cost-effectiveness acceptability curves

Lines illustrate the probability that options are cost-effective at different thresholds for the decision-maker's willingness to pay for SAPY gains

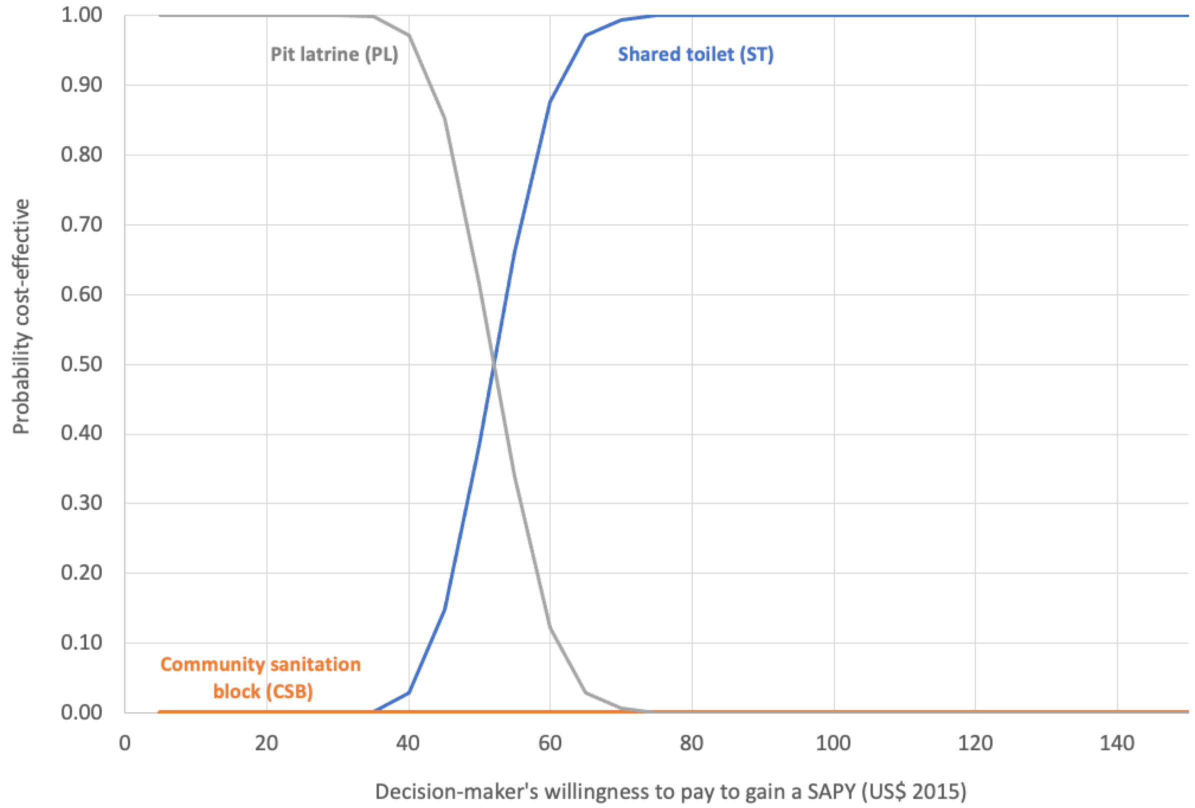
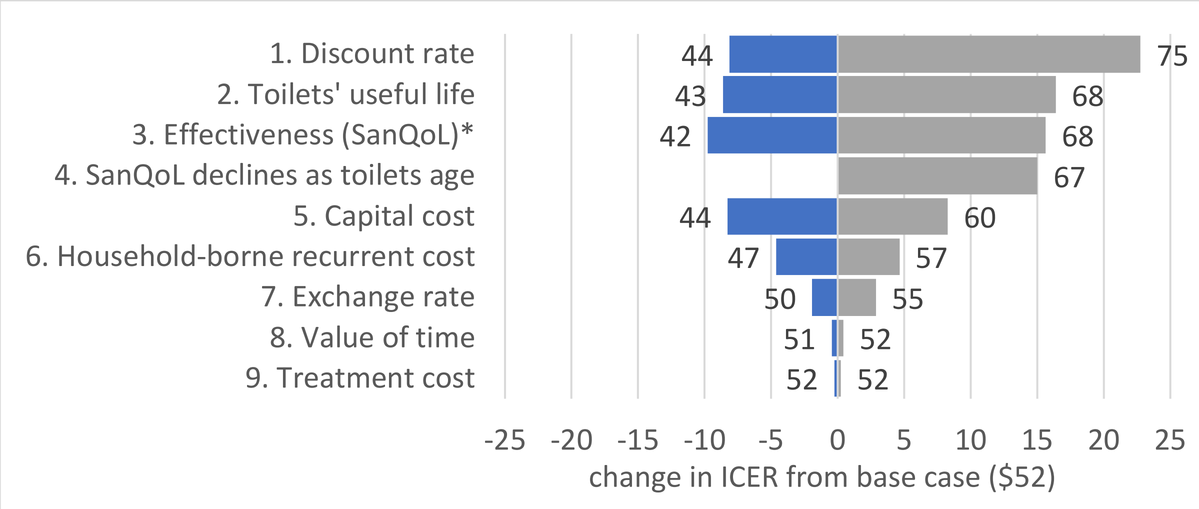


Figure 4 presents a tornado plot indicating the magnitude of changes in the ICER for ST compared to PL when parameters are at high and low plausible values (Supplementary Material B). It also explores changes in other methodological decisions, e.g. whether SanQoL to declines with the life of the toilet. Varying these uncertain methodological assumptions had a relatively small effect on the overall ICER. The discount rate has the largest effect on estimates (bar #1 in Figure 4). When a 10% rate is used instead of the recommended 3% (Wilkinson et al., 2016), ST cost-effectiveness worsens (U$ 75 / SAPY gained – Supplementary Material B). Assumptions about useful life are the next most influential parameter. The base case assumes a 15-year useful life for ST and 5 years for PL. Bar #2 presents a two-way DSA when these are instead 20 and 3 years (lower ICER) or 10 and 7 years respectively (higher ICER). Other scenarios are explained in Supplementary Material B.

Figure 4: Tornado plot of the change in the base-case ICER of ST compared to PL under deterministic sensitivity analysis



Input parameter values for high and low scenarios are reported in Table 2.
*scenarios for the other two comparisons (CSB vs. PL and ST vs. CSB) are in Supplementary Material D. Across all DSA, effectiveness is the only variable for a scenario finds CSB not to be dominated.

Discussion

Globally, 230 million people in urban areas are without "basic" sanitation (UNICEF & WHO, 2019), and poor urban dwellers' ability and willingness to pay for better toilets is often constrained (Tidwell et al., 2019). The allocation of scarce public funds amongst possible urban sanitation interventions is an important policy question. Our cost-effectiveness study evaluates interventions providing heavily subsidised shared toilets (ST) and community sanitation blocks (CSB), as compared to the existing practice of low-quality shared pit latrines (PL) constructed autonomously by households. We present three main findings. First, the economic costs of ST and CSB were seven and 15 times higher than PL, respectively. Second, ST and CSB are both substantially more effective than PL in SAPY terms. Third, ST is cost-effective when the decision-maker's willingness to pay for a SAPY is greater than $52, while CSB is dominated.

The principal strength of our study is its application of a novel approach to valuing sanitation using the sanitation-adjusted person year (SAPY). We worked alongside a trial which concluded that a shared urban sanitation intervention had no impact on under-5 health outcomes. While a DALY-based CEA would be likely to conclude that the intervention was not cost-effective, we demonstrate that a broader value framework based on sanitation-related quality of life brings new insights. A systematic review of economic evaluations (Chapter 4) concluded that no studies had incorporated included quality of life effects of a sanitation intervention. Our SAPY approach thus allows valuation of QoL benefits of sanitation in an economic evaluation for the first time. Furthermore, to our knowledge, no economic evaluation of a public health intervention in a LMIC has previously applied a capability-based outcome. Our results are relevant for urban sanitation decision-makers wanting to know which types of sanitation interventions represent the most efficient use of resources. They are more broadly relevant for the economic evaluation of public health interventions, by illustrating how extra-welfarist measures of value beyond health might be used (Brouwer et al., 2008).

Our study is limited by only including interventions evaluated in the trial, when decision analytic models would ideally consider all possible options for addressing a given problem (Drummond et al., 2015). Future studies could include other interventions such as lower-specification toilets, pay-per-use public toilets, and container-based sanitation. They might also investigate uptake under lower levels of subsidy, since the 85-90% subsidy under the ST and CSB interventions may be unrealistically high for a larger-scale programme. For capital costs of ST and CSB, the retrospective nature of data collection meant relying on NGO expenditure reports and analysing them top-down, which may have introduced bias if the data do not reflect actual resource use (Chapko et al., 2009). For recurrent costs, data collection only three years after ST and CSB construction meant that few households had undertaken emptying and maintenance, which required the introduction of assumptions, e.g. emptying frequency. On the effects side, our estimates were based on an observational study and, while potential confounders were controlled for in regressions, the risk of residual bias remains. There is risk of social desirability bias in any user-reported outcome and, in our study, this could have manifested itself as over-reporting SanQoL in the intervention arm. Similarly, there could have been under-reporting in the control arm out of a desire to receive a future intervention.

The SAPY has some methodological limitations. First, only user-reported QoL is valued, while welfarist CBA can incorporate the value of multiple outcomes. The monetary value of a SAPY could be incorporated in CBA using WTP methods, as with QALYs (Ryen and Svensson, 2015). Second, aspects of sanitation beyond the household may be valuable to society more broadly, such as positive externalities from safe transport, treatment and disposal of faecal waste. However, such benefits are excluded from the SAPY. Third, the SAPY is quite sensitive to assumptions about asset useful life (Figure 4), which are hard to explore empirically. Fourth, our approach to estimating SAPYs assumes that all household members use the same toilet. For adults, this is justified on the basis of our survey data, but younger children might not use toilets. Furthermore, we did not collect SanQoL data for under-18s, who might have different experiences to adults. Lastly, the SanQoL measure may not capture all infrastructure features of CSBs, which had a rooftop water tank and two laundry stations. However, with near-universal piped water in this setting, with limited intermittency (supplementary material D), the incremental benefit of a rooftop tank may be small. Since the SAPY is a novel metric, there is substantial uncertainty about decisions-makers' WTP for SAPY gains. Therefore, it is hard to draw conclusions from the cost-effectiveness acceptability curve, other than the fact that CSB is not cost-effective. For example, when two toilet stalls are needed in a given setting, it would appear to be more efficient to construct two STs than one CSB.

Comparing our cost estimates to cost estimates for similar infrastructure, our survey-based estimate of the financial cost of works ($20) for PL is very similar to the estimate for that technology ($18, converted to 2015 $) from a city-wide survey in Maputo (Bauerl et al., 2016). ST and CSB appear expensive relative to similar infrastructure in other studies. For example, a global costing study estimated that the average financial cost per person for constructing a private flush to septic tank system in urban Sub-Saharan Africa in 2015 was $43 (Hutton and Varughese, 2016), presumably for a household of around five people. The comparable figure in our study is $64 per person for the ST (financial cost of works – Supplementary Material D), noting that these are shared toilets used by an average of 13 people in this setting (Table 4). This implies a cost per toilet three or four times higher than the Hutton and Varughese estimate. Possible explanations for substantially higher costs in this setting are that the ST involved a higher specification design than what a local mason might construct (Mattson, 2016), that construction was contracted to professional companies which might be more expensive than individual masons, and that the cost of business in Mozambique is high relative to other African countries (World Bank, 2019c).

Conclusion

Our results suggest that high-quality shared sanitation offers substantial quality of life gains over low-quality pit latrines. Shared pour-flush toilets are cost-effective if policy-makers willingness to pay to gain a SAPY around $50 or more. This demonstrates the usefulness of the SAPY in assessing quality of life gains, especially in the absence of disease impacts. Such extra-welfarist measures may be useful in other areas of public health (Brouwer et al., 2008). Using the SAPY as an extra-welfarist measure of the value of sanitation could bring insights in other settings where sanitation investment decisions must be made. Further study of the use of SAPYs in CEA could be undertaken in more diverse populations using a wider range of sanitation services. The potential of the SAPY in assessing equity of sanitation interventions could also be explored. However, cost-benefit analysis (CBA) is the most common economic evaluation method for sanitation interventions (Chapter 4). A priority for future research is therefore developing methods for estimating the monetary value of a SAPY, to enable the incorporation of QoL benefits in sanitation CBAs.

Funding. 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 interest. None declared.
Acknowledgements. We acknowledge the hard work of the enumerator teams for the two surveys, without which the study would not have been possible. Furthermore, we greatly appreciate the time and cooperation of respondents who gave their time to be interviewed. We gratefully acknowledge the technical and logistical support received from Vasco Parente and his team at WSUP Mozambique. Thanks are also due to WSUP Mozambique for sharing their cost data, as are staff at the World Bank and CMM. Ian Ross benefitted from discussions with participants at the LSHTM economic evaluation group, the UNC Water and Health Conference 2019, and the UK Health Economics Study Group 2020.

Chapter 10: Discussion and conclusion

10.1. Introduction

Sanitation is the separation of human excreta from human contact (WHO, 2018). Two billion people globally lack access to basic sanitation (UNICEF & WHO, 2019). Improved sanitation can help prevent infectious disease (Cumming et al., 2019; Wolf et al., 2018), but toilet users particularly value other benefits, such as perceived improvements in privacy, safety and dignity (Elmendorf and Buckles, 1980; Jenkins and Curtis, 2005; Solomons, 1978). These are attributes of good quality of life (QoL), and contribute to health in its broadest sense (WHO, 1948).

The sustainable development goal indicator for sanitation (6.2.1) measures "safely managed" sanitation services, the definition of which focuses on toilet quality and treatment of waste (UNICEF & WHO, 2019). While this is appropriate for global monitoring, evaluations of specific interventions require a more nuanced understanding of changes brought about in people's lives, such as the extent to which users' perceptions of privacy have been improved. Quantifying incremental improvements in such outcomes is particularly important for economic evaluation methods such as cost-benefit or cost-effectiveness analysis (Drummond et al., 2015). However, economic evaluations of sanitation interventions have always excluded user-reported outcomes beyond disease prevention and time savings as intangible, in the absence of means for their measurement and valuation (Hutton & Chase, 2016). This may lead to undervaluation of interventions and thus to misallocated resources.

The aim of this thesis was to develop and test an approach to economic evaluation of sanitation interventions by measuring and valuing quality of life. In achieving this aim, my research had four objectives, addressed across the chapters of this thesis: (i) assess the extent to which QoL outcomes have been measured and valued in sanitation economic evaluations (Chapter 4); (ii) to develop a psychometric measure of quality of life related to sanitation (Chapters 5-6); (iii) to estimate the effect of using better toilets on quality of life

(Chapter 7); and (iv) to estimate the cost-effectiveness of a shared sanitation intervention by valuing quality of life gains (Chapters 8-9). This final chapter summarises the main findings across the thesis and reflects on its contribution. In the remainder of this chapter, I discuss the main findings from this thesis by objective (Section 10.2), the limitations of the thesis as a whole (Section 10.3), the overall contribution to knowledge (Section 10.4), and a future research agenda (Section 10.5).

10.2. Main findings

Objective 1: to assess the extent to which QoL outcomes have been measured and valued in sanitation economic evaluations

In Research Paper 1 (chapter 4), I undertook a systematic review which aimed to characterise cost-benefit and cost-effectiveness evidence for drinking water and sanitation interventions in LMICs. Of primary importance for this thesis was the identification of a gap in knowledge concerning how to identify, measure and value quality of life in sanitation economic evaluation. The review demonstrated that user-reported benefits of sanitation such as privacy, safety and dignity, had never yet been included in an economic evaluation of a sanitation intervention. This was diagnosed as being due to the absence of methods for their measurement and valuation, either in monetary or extra-welfarist terms. However, the systematic review made broader contributions. It was the first systematic review of economic evaluations for drinking water and sanitation, and the first review of such studies to systematically assess study quality.

The review also found that welfarist CBA was the dominant economic evaluation method for assessing sanitation interventions. To have impact on policy-makers commissioning economic evaluations and analysts undertaking them, it may therefore be necessary to develop methods for monetary valuation of QoL benefits. While in the latter part of this thesis I have focused on extra-welfarist valuation of sanitation, the design of that work took into account that welfarist valuation may be a possibility. A challenge facing any future work in that direction is that welfarism is at odds with the capability approach, as discussed in Chapter 2. In the systematic review, one reason considered for why economic evaluation is so rarely undertaken in the sanitation sector is that CBA is seen as too burdensome. Therefore, I aimed to make the extra-welfarist valuation methods developed in this thesis compelling enough to encourage alternatives to CBA to be considered.

A further contribution of the systematic review was in identifying that, while drinking water and sanitation interventions in general can deliver net benefits in many settings, benefit-cost ratios are typically modest. This indicates that economic evaluation remains important for guiding resource allocation decisions in the WASH sector. It is therefore all the more important that QoL benefits should be measured and valued since they may tip a finely-balanced analysis in favour of one intervention over another.

In Chapter 2, I identified that health economists, in facing the problem of valuing outcomes beyond health, had often used the capability approach to do so (Sen, 1993). Nine capability-based measures in health economics were identified (Al-Janabi et al., 2012; Canaway et al., 2017; Greco et al., 2015; Grewal et al., 2006; Kinghorn et al., 2014; P. Lorgelly et al., 2015; Netten et al., 2012; Simon et al., 2013; Sutton and Coast, 2014). Chapter 2 also showed that the capability approach has received little attention in the sanitation sector, despite its in potential for assessing sanitation as a state of being. Together, these findings suggested that the capability approach might provide a fruitful means of addressing the knowledge gap around measuring and valuing sanitation-related QoL. Much discourse on sanitation focuses on toilets (commodities) and behaviours (doings). There has been less focus on the value of sanitation in terms of outcomes people experience (beings). The capability approach, with its focus on beings *and* doings, is therefore attractive as a way to conceptualise QoL related to sanitation.

Objective 2: to develop a psychometric measure of quality of life related to sanitation

My qualitative study (Research Paper 2, Chapter 5) aimed to investigate what people most value about sanitation in low-income areas of Maputo, Mozambique, to inform a definition and conceptual model of sanitation-related QoL. Based on in-depth interviews and focus group discussions, five core attributes of sanitation-related QoL were identified: health, disgust, shame, safety and privacy. The paper presents a conceptual model illustrating how sanitation interventions might improve quality of life via changes in these attributes, and how changes are likely to be mediated by conversion factors such as individual and environmental characteristics. This model is the first contribution to knowledge of the qualitative study. I conceptualised sanitation-related QoL as an outcome like health-related QoL, by contrast to insecurity and stress which are usually conceptualised as exposures affecting physical or mental health outcomes. Most qualitative research focused on user-reported consequences of sanitation has been rooted in psychology and public health, so the application of value as an economic lens was also a novel contribution.

From an empirical perspective, it was notable that health (specifically disease risk) was identified as the most-valued attribute of SanQoL in this setting by various different methods. For example, it came out top in the pile-sorting and triad analyses in the qualitative study (Chapter 5) as well as in the ranking exercise in the quantitative survey (Chapter 6). The finding is important because all previous studies assessing relative importance of attributes of sanitation took place in rural areas (Chapter 2), and most had identified health as being in second or third place. As a caveat to this finding, I noted in the discussion sections of Chapters 5 and 6 that some participants had received an intervention, and others had not done so but were likely aware of the intervention in their area, which may have biased their responses in unpredictable ways. In the qualitative study, it was also notable that women and men ranked some attributes differently in the pile-sorting analysis. The biggest absolute differences in scores were for pride, reduced conflict with neighbours (both of which men valued more) and fewer diseases (which women valued more). However, there was no evidence for gendered difference in ranking SanQoL attributes in the quantitative survey (discussed below).

Based on the qualitative findings, Research Paper 3 (Chapter 6) reports a study which aimed to develop and evaluate a novel measure of sanitation-related quality of life (SanQoL). Using a quantitative survey in the MapSan study population, the study followed a stepwise process to develop a five-item SanQoL measure and assess its validity and reliability. There was good evidence in support of 86% of hypothesised presence or absence associations between the five SanQoL attributes and characteristics of the user and their toilet. Further evidence of validity was provided by correlation with the WHO-5 mental wellbeing index and relatively high ICC for two respondents using the same toilet. There was limited evidence for differences in attribute ranking by gender, age and intervention status. Overall, the study provided evidence in support of both validity and reliability of the SanQoL measure.

The contribution of Research Paper 3 was in developing the descriptive system to enable measurement of sanitation-related QoL, as well as in developing a simple ranking method for valuation of sanitation states as an index. The SanQoL measure enables quantification of the extent to which user perceptions of disgust, health risk, shame, safety and privacy have changed as a result of an intervention. Rather than measuring toilet characteristics as most previous measures have done, the items focus on aspects of individuals' QoL in relation to their sanitation practices. This is important because sanitation interventions might improve toilets but with heterogeneous and inequitable impacts on QoL, particularly if appropriate attention is not given to gender and social norms. While the rank sum method is not optimal in that it assumes equal distance between attribute levels, it represents a step forward from equal weighting. SanQoL index values represent the first claim to measuring the value of sanitation to users, as compared to toilet quality measures or the sanitation insecurity profile. The empirical finding that there was no evidence for ranks varying by gender was surprising, as I had expected to find some evidence that women would place more weight on safety and privacy than men. To my knowledge this is also a novel finding, as no other ranking studies that I know of have disaggregated findings by gender.

Objective 3: to estimate the effect of using better toilets on quality of life

Research Paper 4 presented an evaluation study which aimed to estimate the effect of a shared sanitation intervention on quality of life in urban Mozambique. It found strong evidence that using a high-quality pour-flush toilet was associated with a 1.6 standard deviation increase in SanQoL as compared to a low-quality pit latrine. This difference was seen across all five attributes of SanQoL, but with particularly large effect sizes for privacy and safety. There was also good evidence of a 0.2 standard deviation improvement in mental wellbeing as measured by the WHO-5 index. The MapSan trial identified that these interventions did not improve under-5 health outcomes (Knee et al., 2020), but this paper demonstrated that they did improve QoL.

The empirical finding itself is valuable, particularly since the MapSan trial found no impact on infection or disease. The findings on all three outcomes in my study demonstrate that better toilets can improve people's lives beyond reducing infectious disease. The characteristics of the intervention beyond the user interface (e.g. solid walls, lockable door) were likely particularly important in achieving these outcomes. It is particularly relevant that this was a *shared* sanitation intervention, since sanitation facilities shared by more than one household are still considered "limited" in global monitoring. Recall that only 9% of people in Maputo are connected to the sewer and many people live in crowded informal settlements. My results demonstrate that shared sanitation can be high-quality from a QoL perspective, recalling that the modal SanQoL index value in the intervention group was the maximum possible score. Perhaps it was unsurprising that people using better-quality toilets would experience better privacy or less disgust. However, in terms of policy relevance, the bigger contribution was applying two novel ways of quantifying the degree of a QoL difference associated with a given improvement in sanitation service level: SanQoL index values and the sanitation visual analogue scale. This is what enables the subsequent cost-effectiveness study. The WHO-5 finding is also novel in that it is the first time an effect on mental wellbeing has been found in an evaluation of a specific sanitation service level transition as a result of an intervention, rather than assessing many types of services in cross-section.

Objective 4: to estimate the cost-effectiveness of a shared sanitation intervention by valuing quality of life gains

Chapter 8 proposed the sanitation-adjusted person year (SAPY) as an extra-welfarist measure of the value of sanitation to toilet users. To my knowledge, it is the first user-weighted measure of the value of sanitation which combines QoL with person-years experiencing a service. When used in CEA, SAPYs offers a more meaningful way to compare the economic performance of interventions than cost analysis alone. However, this does exclude other benefits beyond quality of life. When used in CBA, SAPYs would allow the summation of QoL benefits along with others, but methods for monetary valuation have not as yet been developed.

The focus on the user experience makes SAPY a measure of outcomes rather than infrastructure or service quality. Everyone defecates and urinates, and within a household, individuals often use the same toilet. However, there can be intra-household variation in SanQoL due to heterogenous experience of the service. The SAPY incorporates the useful life of technologies, allowing time at a given level of SanQoL to be traded off with QoL gains. SAPYs might be used to compare community-led total sanitation to public toilets to hardware subsidies in a given setting. They might also be used to explore equity questions, for example using equity impact planes. In future work, it would be important to consider whether to treat sanitation as a maximand as in QALY-based CEA, or instead to draw on the sufficient capability concept, e.g. comparing "years of sufficient sanitation capability equivalent" rather than SAPYs. For now, though, the SAPY itself represents a step forward in measuring the value of sanitation.

Research Paper 5 applies the SAPY in a cost-effectiveness study, by calculating incremental costs and incremental SAPYs of two sanitation intervention options as compared to existing practice of using low-quality pit latrines. The conclusion of my modelling was that the shared toilet (ST) intervention dominated the community sanitation block (CSB) intervention. The contribution of the study is in demonstrating how SanQoL and SAPYs can be useful in answering empirical questions. Cost-effectiveness planes are familiar to health economists but virtually unknown in the WASH sector, despite the potential they offer in communicating trade-offs and the importance of incremental rather than average analysis. Considering my CEA results themselves, while a DALY-based CEA would be likely to conclude that the intervention was not cost-effective and not be able to distinguish between the two intervention types, the SAPY framework based brought new insights. This is also, to my knowledge, the first time an economic evaluation of a public health intervention in a LMIC has applied a capability-based outcome. The methods and results illustrate how extra-welfarist measures of value beyond health might be used in evaluating public health interventions.

10.3. Limitations

There are important limitations to the research conducted in this thesis as a whole. These limitations relate to three main areas : (i) SanQoL measure reflects the setting, with unknown transferability of the qualitative findings; (ii) valuation using rank sum method could be improved upon; (iii) working alongside a trial; (iv) response bias and previous experience of sanitation.

SanQoL measure reflects the setting, with unknown transferability of the qualitative findings

The qualitative findings (Chapter 5) reflect the setting in which the data were generated, namely a low-income area of urban Maputo where shared sanitation was common. For example, convenience was not raised as an attribute of sanitation-related quality of life, since all respondents already used on-plot toilets. However, in several previous studies, convenience was identified as a valuable attribute of household toilets, as compared to open defecation or public toilets (Novotný et al., 2018). Since the measure was developed on the basis of the qualitative findings (Chapter 6), this may limit the transferability of the measure to other settings, particularly rural areas where the nature of sanitation services and opportunities for open defecation (OD) are different. However, the attributes identified in the qualitative study broadly align with studies in diverse rural and urban areas (Novotný et al., 2018; Sclar et al., 2018), a level of theoretical generalisability may be claimed.

Considering convenience specifically, this is usually already captured in welfarist CBA by valuing time savings, so including it in the SanQoL measure might in any case have opened up issues of double-counting if SAPYs are used in CBA. If the SanQoL measure is used settings with substantially different sanitation practices (e.g. OD, public toilets), further validation of SanQoL through cognitive interviews and psychometric analysis would be recommended. For example, respondents practising OD would need questions reformulated in terms of "carrying out sanitation practices" rather than "using the toilet", which would be straightforward to achieve automatically on mobile-based data collection.

Valuation using rank sum method could be improved upon

The present valuation scheme could be improved upon. Though the rank sum method has previously been used in index valuation (de Kruijk and Rutten, 2007; Greco, 2016), it implicitly assumes equal intervals between response categories. Therefore, it does not strictly speaking meet the Brazier et al. (2016) conditions for use of measures to weight outcomes in economic evaluation. I was unable to apply a trade-off elicitation method as part of this thesis, due to lack of budget to undertake a second survey after the instrument was finalised and validity and reliability investigated. Using the rank sum method in this way is preliminary and illustrative for a first exposition of the SAPY approach. A trade-off study using methods such as discrete choice experiment or best-worst scaling could be undertaken in the future. Using such a value set might change the empirical findings reported in other chapters.

Working alongside a trial

There were substantial benefits to working alongside the MapSan trial (Chapter 4). A specific sanitation intervention had already been delivered with fidelity, and the existence of intervention and control groups would be useful in informing both qualitative and quantitative aspects of the thesis. However, there were also downsides. First, using the trial's allocation groups as the basis for causal inference in Chapter 7 meant relying on the controlled before and after design, which has inherent limitations. In particular, this design precludes randomisation which risks selection bias in the allocation of intervention and control groups. It also necessitates adjusting for covariates, but these may be imprecisely measured or unmeasured, making it impossible to eliminate the risk of unobserved confounding. We relied on MapSan allocation groups in our enrolment strategy and, with SanQoL being a novel measure, it was only possible to collect data after the intervention. This precluded assessment of balance at baseline and adjustment for baseline values in the analysis, and in the four years since the intervention there had been substantial migration into and out of the setting. Finally, use of the intervention and control groups of a trial provided a very narrow set of sanitation service types and living arrangements to inform measure development and assessment of validity and reliability.

Response bias and previous experience of sanitation

Various types of response bias may have affected SanQoL data. For example, there could have been social desirability bias, with those having received the intervention possibly wanting to appear grateful. Those who didn't receive it may have wanted to appear badly off, in the hope of receiving a future intervention. People's subjective valuations of their present states of being are affected by their previous experience, a phenomenon called adaptation (Nussbaum, 2001; Teschl and Comim, 2005). Most people are likely to have experienced something close to "full health" (HRQoL = 1) for some part of their life. However, this is less likely to be the case for SanQoL, perhaps making adaptation a more important consideration. For example, a renter who has moved from dwelling with a high-quality toilet to one with a low-quality one may rate their SanQoL lower than someone who has only ever used a low-quality toilet.

10.4. Contribution to knowledge

Contributions to knowledge were explained across thesis objectives in section 10.2 but are summarised here in terms of methods and empirical findings.

Methods

The development of the two measures used in this thesis, SanQoL (Chapter 6) and the SAPY (Chapter 8), are the most important methodological contributions to knowledge. SanQoL is the first measure to capture quality of life outcomes of sanitation interventions in way that is appropriate for economic evaluation, and the SAPY is the first user-weighted measure of the value of sanitation. Applying the SAPY in an empirical CEA represents a third contribution to methods, since it goes a step beyond simply proposing the SAPY. The systematic review also represents a novel synthesis, since such methods have not been used for economic evaluations studies in WASH before.

There are additional lower-order methodological contributions. The evaluation study (Chapter 7) includes the first use of a visual analogue scale for sanitation, and in the evaluation of a sanitation intervention. The bottom-up costing methods I used for estimating the user-borne recurrent costs of sanitation were, in my view, more rigorous than those used in studies evaluated in the systematic review (Chapter 4). Such bottom-up costing based on assessment of cleaning time, specific cleaning materials and specific maintenance items has not, to my knowledge, been undertaken before.

Empirical findings

The majority of the work in this thesis was focused on developing methods, but several empirical findings also contribute to knowledge. I would highlight four.

First, health (specifically disease risk) was identified as the most-valued attribute of SanQoL in the pile-sorting and triad analyses (Chapter 5) as well as the ranking exercise (Chapter 6). All previous studies assessing relative importance of attributes of sanitation took place in rural areas (Chapter 2), and most had identified health as being in second or third place.

Second, in the measure development study (Chapter 6), an important finding was that there was insufficient evidence for women and men ranking the attributes of SanQoL differently. This is contrary to the supposition that women are likely to place a higher value on privacy and safety. Likewise, in the evaluation study (Chapter 7) there was insufficient evidence that women had benefitted more from the intervention than men. While the sample was powered for differences in SanQoL between intervention and control groups rather than their sub-groups, it is notable that in both cases the p-values were far larger than thresholds generally considered “good” evidence for a difference (e.g. $p<0.05$). The ranking result comprises a comparison of the whole sample by gender, so can be considered a stronger conclusion than the result for lack of gendered benefit from the intervention.

Third, in the evaluation study (Chapter 7), the substantial effect on SanQoL was to some extent was expected, given the leap in quality of service achieved. While it does represent a contribution to knowledge, the more interesting result was perhaps that the intervention improved mental well-being as measured by the WHO-5 index, albeit with a smaller effect size at 0.2 standard deviations. This has only previously been identified in cross-sectional studies, rather than in evaluating a specific intervention.

Fourth, the CEA study (Chapter 9) finding that the ST dominated the CSB was also to some extent unsurprising, given the CSB was an order of magnitude more expensive per person. However, it nonetheless represents an empirical contribution. The systematic review (Chapter 3) identified that most studies, particularly CBAs which are dominant in the sanitation sector, compared one intervention to a “no intervention” scenario. Our study is unusual in that it compared two feasible interventions, and demonstrated the usefulness of economic evaluation in making such comparisons.

10.5. Agenda for future research

A number of avenues for future research have been identified throughout the chapters of this thesis, but five stand out as particularly important.

First, further qualitative and quantitative investigation of sanitation-related quality of life is needed in rural areas and different cultural settings. The nature of sanitation services may be different and, for example, in rural areas there are typically more opportunities for open defecation. While the challenge in urban areas is often improving upon poor-quality existing toilets, in rural areas it is often about getting people to start using toilets in the first place. The comparator in economic evaluation would therefore often be different, and an accurate assessment of the sanitation states of people practising OD is crucial. Further research on the validity and reliability of SanQoL would ideally take place in a larger random sample of a target population using a broader range of sanitation service levels. Related to this, it would be beneficial to include SanQoL in the protocols of impact evaluations in various settings, to begin to establish the effectiveness of different intervention types on quality of life. This may require accompanying qualitative research such as cognitive interviews or think-aloud studies, to ensure the items have face and content validity in those settings. A particular focus of think-aloud work could be on how people scale responses to the questions, rather than the questions themselves.

Second, a specific valuation study for SanQoL would be important in order to fully meet the criteria set out by Brazier et al. (2016). This could employ methods such as a discrete choice experiment or best-worst scaling, to improve upon the rank sum method used presently. This would enable more robust valuation of sanitation states, strengthening the claim of SanQoL truly reflecting the relative value that people place on its different attributes. It would also help communicate to decision-makers how people trade off SanQoL attributes. Further investigation is also required of the properties of the sanitation visual analogue scale, including qualitative work exploring how people approach the exercise.

Third, including QoL benefits in CBA would require a monetary value of the SAPY. A willingness to pay study in an appropriate population would be the first step towards developing methods for this. Such a study could be based on existing methods for estimating the monetary value of QALYs.

Fourth, it would be useful to develop range estimates of SanQoL for given levels of service or toilet types, which could inform future economic evaluations when SanQoL index value data is not available. For example, a study with sampling stratified by types of sanitation, in a population with a broad range of socio-economic status and toilet technologies, would allow estimation of mean SanQoL index values for a broader range of sanitation infrastructure types than those in the MapSan arms. In particular, it would allow the estimation of the incremental effect of different technology characteristics on SanQoL. Assessing the incremental QoL effect of multiple technology and service characteristics would allow identification of those which are associated with the biggest improvements.

Fifth, heterogeneity in SanQoL amongst users with different characteristics requires further exploration, since my study was not powered for sub-group analysis by gender, age, or other relevant characteristics such as disability. Interventions might have differential SanQoL impacts amongst different sub-groups, or even negative impacts for some sub-groups. Knowing this would be relevant in terms of effectiveness and cost-effectiveness.

References

Abelin, C., 1997. The Application of a Cost Benefit Approach on Rural Water Supply Projects. A Minor Field Study of UNICEF's Water Supply Project in Udaipur in Rajahastan. University of Lund.

Acey, C., Kisiangani, J., Ronoh, P., Delaire, C., Makena, E., Norman, G., Levine, D., Khush, R., Peletz, R., 2019. Cross-subsidies for improved sanitation in low income settlements: Assessing the willingness to pay of water utility customers in Kenyan cities. World Dev. 115, 160–177. https://doi.org/https://doi.org/10.1016/j.worlddev.2018.11.006

Agüero, J.M., Bharadwaj, P., 2014. Do the More Educated Know More about Health? Evidence from Schooling and HIV Knowledge in Zimbabwe. Econ. Dev. Cult. Change 62, 489–517. https://doi.org/10.1086/675398

AIAS, 2015. Sanitation and drainage master plan for the greater Maputo Metropolitan Area.

Airoldi, M., Morton, A., 2009. Adjusting life for quality or disability: stylistic difference or substantial dispute? Health Econ. 18, 1237–1247.

Al-Janabi, H., 2018. Do capability and functioning differ? A study of U.K. survey responses. Health Econ. 27, 465–479. https://doi.org/10.1002/hec.3586

Al-Janabi, H., Coast, J., Flynn, T.N., 2008. What do people value when they provide unpaid care for an older person? A meta-ethnography with interview follow-up. Soc. Sci. Med. 67, 111–121. https://doi.org/10.1016/j.socscimed.2008.03.032

Al-Janabi, H., Flynn, T.N., Coast, J., 2012. Development of a self-report measure of capability wellbeing for adults: The ICECAP-A. Qual. Life Res. 21, 167–176. https://doi.org/10.1007/s11136-011-9927-2

Andrés, L., Briceño, B., Chase, C., Echenique, J.A., 2017. Sanitation and externalities: evidence from early childhood health in rural India. J. Water Sanit. Hyg. Dev. 7, 272–289. https://doi.org/10.2166/washdev.2017.143

Angrist, N., Evans, D., Filmer, D., Glennerster, R., Rogers, H., Sabarwal, S., 2020. How to Improve Education Outcomes Most Efficiently? A Comparison of 150 Interventions Using the New Learning-Adjusted Years of Schooling Metric (No. 9450), Policy Research Working Papers.

Aunger, R., Curtis, V., 2013. The Anatomy of Motivation: An Evolutionary-Ecological Approach. Biol. Theory 8, 49–63. https://doi.org/10.1007/s13752-013-0101-7

Baltussen, R., Niessen, L., 2006. Priority setting of health interventions: The need for multi-criteria decision analysis. Cost Eff. Resour. Alloc. 4, 1–9. https://doi.org/10.1186/1478-7547-4-14

Barrington, D., Sridharan, S., Saunders, S.G., Amjad, U.Q., Bartram, J., Shields, K., Souter, R., Meo, S., Aalbersberg, W., 2014. Embedding a capability approach within sanitation marketing, in: UNC Water and Health 2014 Conference: Posters. Chapel Hill, North Carolina, U.S.A.

Barrington, D., Sridharan, S., Shields, K.F., Saunders, S.G., Souter, R.T., Bartram, J., 2017. Sanitation marketing: A systematic review and theoretical critique using the capability approach. Soc. Sci. Med. 194, 128–134. https://doi.org/10.1016/j.socscimed.2017.10.021

Bartram, J., Cairncross, S., 2010. Hygiene, sanitation, and water: forgotten foundations of health. PLoS Med. 7, e1000367. https://doi.org/http://dx.doi.org/10.1371/journal.pmed.1000367

Barungi, M., Kasirye, I., 2011. Cost-effectiveness of water interventions: the case for public

stand-posts and bore-holes in reducing diarrhoea among urban households in Uganda. Res. Ser. - Econ. Policy Res. Cent. 16-pp.
Batley, R., Mcloughlin, C., 2015. The Politics of Public Services: A Service Characteristics Approach. World Dev. 74, 275–285. https://doi.org/10.1016/j.worlddev.2015.05.018
Bauerl, M., Arsénio, A.M., Muximpua, O., Hawkins, P., 2016. Emptying of sanitation facilities in Maputo, Mozambique. Assessment of Attitudes and Practices at a Household Level.
Bhalotra, S.R., Diaz-Cayeros, A., Miller, G., Miranda, A., Venkataramani, A.S., 2017. Urban water disinfection and mortality decline in developing countries. Work. Pap. - Cent. Glob. Dev. 44-pp.
Bilinski, A., Neumann, P., Cohen, J., Thorat, T., McDaniel, K., Salomon, J.A., 2017. When cost-effective interventions are unaffordable: Integrating cost-effectiveness and budget impact in priority setting for global health programs. PLOS Med. 14, e1002397. https://doi.org/10.1371/journal.pmed.1002397
Bisung, E., Elliott, S.J., 2017. Psychosocial impacts of the lack of access to water and sanitation in low- and middle-income countries: A scoping review. J. Water Health 15, 17–30. https://doi.org/10.2166/wh.2016.158
Bleichrodt, H., Pinto, J.L., 2005. The validity of QALYS under non-expected utility. Econ. J. 115, 533–550. https://doi.org/10.1111/j.1468-0297.2005.00999.x
Boardman, A., Greenberg, D., Vining, A., Weimer, D., 2018. Cost-Benefit Analysis: Concepts and Practices, 5th ed. Cambridge University Press.
Bobinac, A., Van Exel, N.J.A., Rutten, F.F.H., Brouwer, W.B.F., 2010. Willingness to pay for a quality-adjusted life-year: The individual perspective. Value Heal. 13, 1046–1055. https://doi.org/10.1111/j.1524-4733.2010.00781.x
Boije, E., Eskhult, G., 2014. With a few liters of clean drinking water- A Cost Benefit Analysis of the socio-economic effects from implementing new clean drinking water technologies in rural India. Swedish University of Agricultural Sciences.
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 Plan. 17, 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.
Brazier, J., Rowen, D., Karimi, M., Peasgood, T., Tsuchiya, A., Ratcliffe, J., 2018. Experience-based utility and own health state valuation for a health state classification system: why and how to do it. Eur. J. Health Econ. 19, 881–891. https://doi.org/10.1007/s10198-017-0931-5
Briggs, A.H., Weinstein, M.C., Fenwick, E.A.L., Karnon, J., Sculpher, M.J., Paltiel, A.D., 2012. Model parameter estimation and uncertainty analysis: A report of the ISPOR-SMDM modeling good research practices task force working group-6. Med. Decis. Mak. 32, 722–732. https://doi.org/10.1177/0272989X12458348
Briscoe, J., 1987. A role for water supply and sanitation in the child survival revolution. Bull. Pan Am. Health Organ. 21, 93–105.
Briscoe, J., 1984. Water supply and health in developing countries: selective primary health care revisited. Am. J. Public Health 74, 1009–1013.
Broome, J., 2004. Weighing Lives. Oxford University Press, Oxford.
Brouwer, W.B.F., Culyer, A.J., van Exel, N.J.A., Rutten, F.F.H., 2008. Welfarism vs. extra-welfarism. J. Health Econ. 27, 325–338. https://doi.org/10.1016/j.jhealeco.2007.07.003

Brown, J., Cumming, O., Bartram, J., Cairncross, S., Ensink, J., Holcomb, D., Knee, J., Kolsky, P., Liang, K., Liang, S., Nala, R., Norman, G., Rheingans, R., Stewart, J., Zavale, O., Zuin, V., 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, 1–12. https://doi.org/10.1136/bmjopen-2015-008215

Burt, Z., Ercumen, A., Billava, N., Ray, I., 2018. From intermittent to continuous service: Costs, benefits, equity and sustainability of water system reforms in Hubli-Dharwad, India. WORLD Dev. 109, 121–133. https://doi.org/10.1016/j.worlddev.2018.04.011

Cairncross, S., Valdmanis, V., 2006. Water supply, sanitation, and hygiene promotion. Dis. Control priorities Dev. Ctries. 771–792.

Canaway, A., Al-janabi, H., Kinghorn, P., Bailey, C., Coast, J., 2017. Development of a measure (ICECAP-Close Person Measure) through qualitative methods to capture the benefits of end-of-life care to those close to the dying for use in economic evaluation. https://doi.org/10.1177/0269216316650616

Capone, D., Buxton, H., Cumming, O., Dreibelbis, R., Knee, J., Nalá, R., Ross, I., Brown, J., 2020. Impact of an intervention to improve pit latrine emptying practices in low income urban neighborhoods of Maputo, Mozambique. Int. J. Hyg. Environ. Health 226, 113480. https://doi.org/10.1016/j.ijheh.2020.113480

Caruso, B., Sinharoy, S., Conrad, A., Patrick, M., 2020. Briefing note - Measuring Urban Sanitation and Empowerment (MUSE).

Caruso, B.A., Clasen, T., Yount, K.M., Cooper, H.L.F., Hadley, C., Haardörfer, R., 2017a. Assessing women's negative sanitation experiences and concerns: The development of a novel sanitation insecurity measure. Int. J. Environ. Res. Public Health 14, 1–22. https://doi.org/10.3390/ijerph14070755

Caruso, B.A., Clasen, T.F., Hadley, C., Yount, K.M., Haardörfer, R., Rout, M., Dasmohapatra, M., Cooper, H.L., 2017b. Understanding and defining sanitation insecurity: women's gendered experiences of urination, defecation and menstruation in rural Odisha, India. BMJ Glob. Heal. 2, e000414. https://doi.org/10.1136/bmjgh-2017-000414

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 - Popul. Heal. 5, 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.

Cha, S., Cho, Y., Kim, S.J., Lee, Y.J., Choi, S., Asuming, P., Kim, Y., Jin, Y., 2018. Cost-benefit analysis of water source improvements through borehole drilling or rehabilitation: an empirical study based on a cluster randomized controlled trial in the Volta Region, Ghana. Glob. Health Action 11. https://doi.org/10.1080/16549716.2018.1523303

Cha, S., Jung, S., Bizuneh, D.B., Abera, T., Doh, Y.A., Seong, J., Ross, I., 2020. Benefits and costs of a community-led total sanitation intervention in rural Ethiopia—A trial-based ex post economic evaluation. Int. J. Environ. Res. Public Health 17, 1–21. https://doi.org/10.3390/ijerph17145068

Chapko, M.K., Liu, C.-F., Perkins, M., Li, Y.-F., Fortney, J.C., Maciejewski, M.L., 2009. Equivalence of two healthcare costing methods: bottom-up and top-down. Health Econ. 18, 1188–1201. https://doi.org/10.1002/hec.1422

Chuan, L., Hutton, G., Liqiong, Y., Jinmin, F., Tiwei, Z., Lin, D., Pu, Z., Ronghuai, L., 2012.

Economic assessment of sanitation interventions in Yunnan Province, People's Republic of China. World Bank.
Churchill, A., de Ferranti, D., Roche, R., Tager, C., Walters, A., Yazer, A., 1987. Rural water supply and sanitation: Time for a change, World Bank Discussion Papers.
Cicchetti, D. V, 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol. Assess. https://doi.org/10.1037/1040-3590.6.4.284
Clasen, T., Haller, L., Walker, D., Bartram, J., Cairncross, S., 2007. Cost-effectiveness of water quality interventions for preventing diarrhoeal disease in developing countries. J. Water Health 5, 599–608. https://doi.org/10.2166/wh.2007.010
Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P.C., Sculpher, M., 2015. Methods for the estimation of the National Institute for Health and care excellence cost-effectiveness threshold. Health Technol. Assess. (Rockv). 19, 1–503. https://doi.org/10.3310/hta19140
CMM, 2018. Orçamento aprovado para 2017. Direcçao Municipal de Infra-estrutura.
CMM, 2017. Postura de Saneamento e Drenagem. Assembleia Municipal de Maputo, Mozambique.
Coast, J., 2019. Assessing capability in economic evaluation: a life course approach? Eur. J. Heal. Econ. 20, 779–784. https://doi.org/10.1007/s10198-018-1027-6
Coast, J., 2009. Maximisation in extra-welfarism: A critique of the current position in health economics. Soc. Sci. Med. 69, 786–792. https://doi.org/10.1016/j.socscimed.2009.06.026
Coast, J., 2004. Is economic evaluation in touch with society's health values? Br. Med. J. 329, 1233–1236. https://doi.org/10.1136/bmj.329.7476.1233
Coast, J., Flynn, T.N., Natarajan, L., Sproston, K., Lewis, J., Louviere, J.J., Peters, T.J., 2008a. Valuing the ICECAP capability index for older people. Soc. Sci. Med. 67, 874–882. https://doi.org/10.1016/j.socscimed.2008.05.015
Coast, J., Jackson, L., 2017. Understanding primary data analysis, in: Qualitative Methods for Health Economics. Rowman & Littlefield, pp. 92–118.
Coast, J., Kinghorn, P., Mitchell, P., 2015. The Development of Capability Measures in Health Economics: Opportunities, Challenges and Progress. Patient 8, 119–126. https://doi.org/10.1007/s40271-014-0080-1
Coast, J., Kinghorn, P., Owen-Smith, A., 2017. Designing qualitative research studies in health economics, in: Qualitative Methods for Health Economics. Rowman & Littlefield, pp. 19–41.
Coast, J., Smith, R., Lorgelly, P., 2008b. Should the capability approach be applied in Health Economics? Health Econ. 17, 667–670.
Coast, J., Smith, R.D., Lorgelly, P., 2008c. Welfarism, extra-welfarism and capability: The spread of ideas in health economics. Soc. Sci. Med. 67, 1190–1198. https://doi.org/10.1016/j.socscimed.2008.06.027
Coffey, D., Gupta, A., Hathi, P., Khurana, N., Spears, D., Srivastav, N., Vyas, S., 2014. Revealed preference for open defecation: Evidence from a new survey in rural north India. Econ. Polit. Wkly. 49, 43–55.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Routledge.
Connell, J., Carlton, J., Peasgood, T., Mukuria, C., Brazier, J., Augustovski, F., Belizan, M., Engel, L., Greiner, W., Hanna, L., Ludwig, K., Monteiro, A., Mulhern, B., Nan, L., Pickard, S., 2018. Extending the QALY: developing and testing the proposed items for a new

generic measure - results from qualitative review and face validity with patients, social care users and carers. Qual. Life Res. 27, S189.
Cookson, R., 2005. QALYs and the capability approach. Health Econ. 14, 817–829. https://doi.org/10.1002/hec.975
Cookson, R., Mirelman, A.J., Griffin, S., Asaria, M., Dawkins, B., Norheim, O.F., Verguet, S., J Culyer, A., 2017. Using Cost-Effectiveness Analysis to Address Health Equity Concerns. Value Health 20, 206–212. https://doi.org/10.1016/j.jval.2016.11.027
Copenhagen Consensus, 2020. Copenhagen Consensus Projects [WWW Document]. URL https://www.copenhagenconsensus.com/projects (accessed 9.4.20).
Corbin, J., Strauss, A., 1990. Grounded Theory Research : Procedures , Canons , and Evaluative Criteria. Qual. Sociol. 13, 3–21.
Crocker, J., Saywell, D., Bartram, J., 2017. Sustainability of community-led total sanitation outcomes: Evidence from Ethiopia and Ghana. Int. J. Hyg. Environ. Health 220, 551–557. https://doi.org/10.1016/j.ijheh.2017.02.011
Cronin, A.A., Ohikata, M., Kumar, M., 2014. Social and economic cost-benefit analysis of sanitation in Odisha State, India. J. WATER Sanit. Hyg. Dev. 4, 521–531. https://doi.org/10.2166/washdev.2014.150
Culyer, A.J., 1989. The normative economics of health care finance and provision. Oxford Rev. Econ. Policy 5, 34–56.
Cumming, O., Arnold, B.F., Ban, R., Clasen, T., Esteves Mills, J., Freeman, M.C., Gordon, B., Guiteras, R.P., Howard, G., Hunter, P.R., Johnston, R.B., Pickering, A.J., Prendergast, A.J., Prüss-Ustün, A., Rosenboom, J.W., Spears, D., Sundberg, S., Wolf, J., Null, C., Luby, S.P., Humphrey, J.H., Colford, J.M., 2019. The implications of three major new trials for the effect of water, sanitation and hygiene on childhood diarrhea and stunting: a consensus statement. BMC Med. 17, 173. https://doi.org/10.1186/s12916-019-1410-x
Cvjetanovic, B., Grab, B., 1976. Rough determination of the cost benefit balance point of sanitation programmes. Bull. World Health Organ. 54, 207–215.
Das, D.K., Islam, M.S., Hadiujjaman, S., Dutta, C.B., Morshed, M.M., 2019. Health cost of salinity contamination in drinking water: evidence from Bangladesh. Environ. Econ. POLICY Stud. 21, 371–397. https://doi.org/10.1007/s10018-018-0234-9
Dasgupta, P., Sahay, S., Prakash, A., Lutz, A., 2019. Cost effective adaptation to flood: sanitation interventions in the Gandak river basin, India. Clim. Dev. https://doi.org/10.1080/17565529.2019.1682490
de Kruijk, H., Rutten, M., 2007. Weighting dimensions of poverty based on peoples priorities: Constructing a Composite Poverty Index for the Maldives.
Delea, M.G., Snyder, J.S., Belew, M., Caruso, B.A., Garn, J. V., Sclar, G.D., Woreta, M., Zewudie, K., Gebremariam, A., 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, 1–15. https://doi.org/10.1186/s12889-019-7040-6
Devleesschauwer, B., Havelaar, A.H., Maertens De Noordhout, C., Haagsma, J.A., Praet, N., Dorny, P., Duchateau, L., Torgerson, P.R., Van Oyen, H., Speybroeck, N., 2014. DALY calculation in practice: A stepwise approach. Int. J. Public Health 59, 571–574. https://doi.org/10.1007/s00038-014-0553-y
Devlin, N.J., Shah, K.K., Feng, Y., Mulhern, B., van Hout, B., 2018. Valuing health-related quality of life: An EQ-5D-5L value set for England. Heal. Econ. (United Kingdom) 27, 7–

22. https://doi.org/10.1002/hec.3564
Devlin, N.J., Sussex, J., 2011. Incorporating Multiple Criteria in HTA: Methods and Processes, Office of Health Economics.
Dickin, S., Bisung, E., Nansi, J., Charles, K., 2021. Empowerment in water, sanitation and hygiene index. World Dev. 137, 105158. https://doi.org/10.1016/j.worlddev.2020.105158
Dickinson, K.L., Patil, S.R., Pattanayak, S.K., Poulos, C., Yang, J.H., 2015. Nature's call: impacts of sanitation choices in Orissa, India. Econ. Dev. Cult. Change 64, 1–29. https://doi.org/http://dx.doi.org/10.1086/682958
Doyal, L., Gough, I., 1984. A theory of human needs. Crit. Soc. Policy 4, 6–38. https://doi.org/10.1177/026101838400401002
Dreibelbis, R., Jenkins, M., Chase, R.P., Torondel, B., Routray, P., Boisson, S., Clasen, T., Freeman, M.C., 2015. Development of A Multidimensional Scale to Assess Attitudinal Determinants of Sanitation Uptake and Use. Environ. Sci. Technol. 49, 13613–13621. https://doi.org/10.1021/acs.est.5b02985
Drummond, M., Stoddard, G.L., Torrance, G.W., 2015. Methods for the economic evaluation of health care programmes, fourth. ed. Oxford University Press, Oxford, UK.
Dupas, P., Nhlema, B., Wagner, Z., Wolf, A., Wroe, E., 2020. Expanding Access to Clean Water for the Rural Poor: Experimental Evidence from Malawi (No. 27570), NBER working papers.
Dwumfour-Asare, B., Larsen, B., Wong, B., 2020. Benefit-cost analysis of improved urban sanitation in Ghana. Copenhagen Consensus - Ghana Priorities.
Edwards, J.A., Van Laar, D., Easton, S., Kinman, G., 2009. The Work-related Quality of Life Scale for Higher Education Employees. Qual. High. Educ. 15, 207–219. https://doi.org/10.1080/13538320903343057
Eklund, S., Herrmann, T., 1991. The application of cost-benefit analysis on rural water supply projects. Växjö University.
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, Washington, D.C.
Elsevier, 2019. Mendeley Desktop.
Encyclopædia Britannica, 2019. Maputo: national capital, Mozambique. Encycl. Br.
Enssle, S., 2010. Promoting Ecosan in Developing Countries – A Case Study in Tamil Nadu, India – Based on the Capability Approach. University of Basel - MSc thesis.
Euroqol Group, 2009. EuroQol Group EQ-5D™ Health Questionnaire. Http://Www.Euroqol.Org/Eq-5D-Products/How-To-Obtain-Eq-5D.Html 3.
Evans, D.K., Popova, A., 2016. Cost-Effectiveness Analysis in Development: Accounting for Local Costs and Noisy Impacts. World Dev. 77, 262–276. https://doi.org/10.1016/j.worlddev.2015.08.020
Fahimuddin, 2012. Drinking water collection and cost-benefit analysis of a rural water supply scheme in Uttarakhand State. J. Rural Dev. 31, 1–15.
Fayers, P., Machin, D., 2015. Quality of Life: The Assessment, Analysis and Interpretation of Patient-Reported Outcomes, 3rd ed. Wiley.
Feachem, R.G., 1986. Preventing diarrhoea: what are the policy options? Health Policy Plan. 1, 109–117. https://doi.org/10.1093/heapol/1.2.109
Filmer, D., Rogers, H., Angrist, N., Sabarwal, S., 2020. Learning-adjusted years of schooling (LAYS): Defining a new macro measure of education. Econ. Educ. Rev.

https://doi.org/10.1016/j.econedurev.2020.101971
Flynn, T.N., Louviere, J.J., Peters, T.J., Coast, J., 2007. Best – worst scaling : What it can do for health care research and how to do it 26, 171–189. https://doi.org/10.1016/j.jhealeco.2006.04.002
Franklin, M., Payne, K., Elliott, R.A., 2018. Quantifying the Relationship between Capability and Health in Older People: Can't Map, Won't Map. Med. Decis. Mak. 38, 79–94. https://doi.org/10.1177/0272989X17732975
Garn, J. V, Sclar, G.D., Freeman, M.C., Penakalapati, G., Alexander, K.T., Brooks, P., Rehfuess, E.A., Boisson, S., Medlicott, K.O., Clasen, T.F., 2017. The impact of sanitation interventions on latrine coverage and latrine use : A systematic review and meta-analysis. Int. J. Hyg. Environ. Health 220, 329–340. https://doi.org/10.1016/j.ijheh.2016.10.001
GBD Network, 2018. Global Burden of Disease Study 2017 (GBD 2017) Disability Weights [WWW Document]. Inst. Heal. Metrics Eval. URL http://ghdx.healthdata.org/record/ihme-data/gbd-2017-disability-weights (accessed 8.17.20).
Gold, M.R., Stevenson, D., Fryback, D.G., 2002. HALYs and QALYs and DALYs, Oh My: Similarities and Differences in Summary Measures of Population Health. Annu. Rev. Public Health 23, 115–134. https://doi.org/10.1146/annurev.publhealth.23.100901.140513
Gomersall, J.S., Jadotte, Y.T., Xue, Y., Lockwood, S., Riddle, D., Preda, A., 2015. Conducting systematic reviews of economic evaluations. Int. J. Evid. Based. Healthc. 13, 170–178. https://doi.org/10.1097/XEB.0000000000000063
González, A.H., Aristizábal, A.B., 2011. Community-Led Total Sanitation: A Technologies for Freedom approach to sanitation?, in: Human Development and Capabilities Association Conference 2011.
Goranitis, I., Coast, J., Day, E., Copello, A., Freemantle, N., Frew, E., 2017. Maximizing Health or Sufficient Capability in Economic Evaluation? A Methodological Experiment of Treatment for Drug Addiction. Med. Decis. Mak. 37, 498–511. https://doi.org/10.1177/0272989X16678844
Greco, G., 2016. Setting the Weights: The Women's Capabilities Index for Malawi. Soc. Indic. Res. 135, 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 Econ. 25, 83–94. https://doi.org/10.1002/hec.3302
Greco, G., Skordis-Worrall, J., Mkandawire, B., Mills, A., 2015. What is a good life? Selecting capabilities to assess women's quality of life in rural Malawi. Soc. Sci. Med. 130, 69–78. https://doi.org/10.1016/j.socscimed.2015.01.042
Grewal, I., Lewis, J., Flynn, T., Brown, J., Bond, J., Coast, J., 2006. Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? Soc. Sci. Med. 62, 1891–1901. https://doi.org/10.1016/j.socscimed.2005.08.023
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. J. Water Health 9, 617–627. https://doi.org/10.2166/wh.2011.198
Gross, E., Günther, I., 2014. Why do households invest in sanitation in rural Benin: Health, wealth, or prestige? Water Resour. Res. 50, 8314–8329. https://doi.org/doi:10.1002/

2014WR015899
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, 177. https://doi.org/10.1186/1471-2458-12-177
Gunther, I., Fink, G., 2013. Saving a life-year and reaching MDG 4 with investments in water and sanitation: a cost-effective policy? Eur. J. Dev. Res. 25, 129–153. https://doi.org/http://dx.doi.org/10.1057/ejdr.2012.24
Gyrd-Hansen, D., 2003. Willingness to pay for a QALY. Health Econ. 12, 1049–1060. https://doi.org/10.1002/hec.799
Haller, L., Hutton, G., Bartram, J., 2007. Estimating the costs and health benefits of water and sanitation improvements at global level, Journal of Water and Health. https://doi.org/10.2166/wh.2007.008
Hawker, G.A., Mian, S., Kendzerska, T., French, M., 2011. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF. Arthritis Care Res. (Hoboken). 63, S240–S252. https://doi.org/10.1002/acr.20543
Hawkins, P., Muximpua, O., 2015. Developing Business Models for Fecal Sludge Management in Maputo. World Bank.
Heng, S.S., Hutton, G., Kongchheng, P., Kov, P., 2012. Economic Assessment of Sanitation Interventions in Cambodia.
Hensher, D., Rose, J., Greene, W., 2015. Applied Choice Analysis, 2nd ed. Cambridge University Press.
Hoch, J.S., Dewa, C.S., 2008. A clinician's guide to correct cost-effectiveness analysis: Think incremental not average. Can. J. Psychiatry 53, 267–274. https://doi.org/10.1177/070674370805300408
Hope, S.F., Webster, J., Trieu, K., Pillay, A., Ieremia, M., Bell, C., Snowdon, W., Neal, B., Moodie, M., 2017. A systematic review of economic evaluations of population-based sodium reduction interventions. PLoS One 12, e0173600.
Hulland, K., Chase, R.P., Caruso, B.A., Swain, R., Biswal, B., Sahoo, K.C., Panigrahi, P., Dreibelbis, R., 2015. Sanitation, stress, and life stage: A systematic data collection study among women in Odisha, India. PLoS One 10. https://doi.org/10.1371/journal.pone.0141883
Humphrey, J.H., Mbuya, M.N.N., Ntozini, R., ..., Makoni, T., 2019. Independent and combined effects of improved water, sanitation, and hygiene, and improved complementary feeding, on child stunting and anaemia in rural Zimbabwe: a cluster-randomised trial. Lancet Glob. Heal. 7, e132–e147. https://doi.org/10.1016/S2214-109X(18)30374-7
Hunter, R., Shearer, J., 2014. Cost-consequences analysis - an underused method of economic evaluation, National Institute for Health Research.
Hurley, J., 2000. An overview of the normative economics of the health care sector, in: Culyer, A.J., Newhouse, J. (Eds.), Handbook of Health Economics. Elsevier.
Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., Augustovski, F., Briggs, A.H., Mauskopf, J., Loder, E., 2013. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ Br. Med. J. 346, f1049. https://doi.org/10.1136/bmj.f1049
Hutton, G., 2015. Benefits and Costs of the Water Sanitation and Hygiene Targets for the

Post-2015 Development Agenda, Post-2015 Consensus. Copenhagen Consensus.
Hutton, G., 2013. Global costs and benefits of reaching universal coverage of sanitation and drinking-water supply. J. Water Health 11, 1–12. https://doi.org/10.2166/wh.2012.105
Hutton, G., Chase, C., 2016. The knowledge base for achieving the sustainable development goal targets on water supply, sanitation and hygiene. Int. J. Environ. Res. Public Health 13. https://doi.org/10.3390/ijerph13060536
Hutton, G., Haller, L., 2004. Evaluation of the Costs and Benefits of Water and Sanitation Improvements at the Global Level, World Health Organization. https://doi.org/10.2166/wh.2007.008
Hutton, G., Haller, L., Bartram, J., 2007a. Global cost-benefit analysis of water supply and sanitation interventions. J. Water Health 5, 481–501. https://doi.org/10.2166/wh.2007.009
Hutton, G., Haller, L., Bartram, J., 2007b. Economic and health effects of increasing coverage of low cost household drinking-water supply and sanitation interventions to countries off-track to meet MDG target 10. World Health Organisation.
Hutton, G., Odhiambo, F., Osbert, N., Kumar, A., Patil, S., 2017. Financial and Economic Impacts of the Swachh Bharat Mission in India. New Delhi, India.
Hutton, G., Patil, S., Kumar, A., Osbert, N., Odhiambo, F., 2020. Comparison of the costs and benefits of the Clean India Mission. World Dev. 134. https://doi.org/10.1016/j.worlddev.2020.105052
Hutton, G., Rodriguez, U.P., Winara, A., Viet-Anh, N., Phyrum, K., Chuan, L., Blackett, I., Weitz, A., 2014. Economic efficiency of sanitation interventions in Southeast Asia. J. Water Sanit. Hyg. Dev. 4, 23–36. https://doi.org/10.2166/washdev.2013.158
Hutton, G., Varughese, M., 2016. The Costs of Meeting the 2030 Sustainable Development Goal Targets on Drinking Water, Sanitation, and Hygiene: Summary Report. Costs Meet. 2030 Sustain. Dev. Goal Targets Drink. Water, Sanit. Hyg. Summ. Rep. https://doi.org/10.1596/k8632
Hutubessy, R., Chisholm, D., Tan-Torres Edejer, T., 2003. Cost Effectiveness and Resource Allocation Generalized cost-effectiveness analysis for national-level priority-setting in the health sector Address: 1 Stop TB Programme (STB), HIV/AIDS, TB and Malaria cluster (HTM), World Health Organization and. Cost Eff. Resour. Alloc. 1, 1–13. https://doi.org/10.1186/1478-7547-1-8
Ignacio, J., 2006. A Comparative Benefit-Cost and Cost-Effectiveness Analysis of Ecosan and Non-Ecosan Systems in San Fernando City.
Immler, U., 2018. Sanitation Realities in Peri-Urban Communities : Unfreedoms , Capabilities and the Conscious Mind - A Case of Chennai , India. University of Bradford - PhD thesis.
INE, 2019. IV Recenseamento Geral da População e Habitação, 2017: Resultados Definitivos, Instituto Nacional de Estatística, Maputo-Moçambique.
International Rescue Committee, 2016. Cost Efficiency Analysis: Latrine-Building Programs in Ethiopia.
Investing.com, 2019. USD/MZN - US Dollar Mozambican Metical [WWW Document]. URL https://uk.investing.com/currencies/usd-mzn-historical-data (accessed 12.2.19).
Jain, A., Subramanian, S. V., 2018. Intrinsic and instrumental perspectives to sanitation. SSM - Popul. Heal. 5, 267–269. https://doi.org/10.1016/j.ssmph.2018.07.005
Jenkins, M., 1999. Sanitation Promotion in Developing Countries: Why the Latrines of Benin are Few and Far Between. University of California Davis. https://doi.org/10.1007/s13398-014-0173-7.2

Jenkins, M., Curtis, V., 2005. Achieving the “good life”: Why some people want latrines in rural Benin. Soc. Sci. Med. 61, 2446–2459. https://doi.org/10.1016/j.socscimed.2005.04.036

Jenkins, M., Scott, B., 2007. Behavioral indicators of household decision-making and demand for sanitation and potential gains from social marketing in Ghana. Soc. Sci. Med. 64, 2427–2442. https://doi.org/10.1016/J.SOCSCIMED.2007.03.010

Jenkins, M., Sugden, S., 2006. Rethinking Sanitation: Lessons and Innovation for Sustainability and Success in the New Millennium. Human Development Occasional Papers (1992–2007) HDOCPA-2006–27.

Jeuland, M., Whittington, D., 2009. Cost-benefit comparisons of investments in improved water supply and cholera vaccination programs. Vaccine 27, 3109–3120. https://doi.org/10.1016/j.vaccine.2009.02.104

Jeuland, M., Whittington, D., M., J., 2009. Cost-benefit comparisons of investments in improved water supply and cholera vaccination programs. Vaccine 27, 3109–3120. https://doi.org/http://dx.doi.org/10.1016/j.vaccine.2009.02.104

Jeuland, M.A., Fuente, D.E., Ozdemir, S., Allaire, M.C., Whittington, D., 2013. The Long-Term Dynamics of Mortality Benefits from Improved Water and Sanitation in Less Developed Countries. PLoS One 8. https://doi.org/10.1371/journal.pone.0074804

Jha, P., Bangoura, O., Ranson, K., 1998. The Cost-Effectiveness of Forty Health Interventions in Guinea . Health Policy Plan. 13, 249–262.

JICA, 2014. Comprehensive Urban Transport Master Plan for the Greater Maputo Area.

Karimi, M., Brazier, J., 2016. Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? Pharmacoeconomics 34, 645–649. https://doi.org/10.1007/s40273-016-0389-9

King, C.H., 2014. Health metrics for helminth infections. Acta Trop. 141, (in press; doi: 10.1016/j.actatropica.2013.12.001). https://doi.org/10.1016/j.actatropica.2013.12.001

Kinghorn, P., 2019. Using deliberative methods to establish a sufficient state of capability well-being for use in decision-making in the contexts of public health and social care. Soc. Sci. Med. 240, 112546. https://doi.org/https://doi.org/10.1016/j.socscimed.2019.112546

Kinghorn, P., Robinson, A., Smith, R.D., 2014. Developing a Capability-Based Questionnaire for Assessing Well-Being in Patients with Chronic Pain. Soc. Indic. Res. 120, 897–916. https://doi.org/10.1007/s11205-014-0625-7

Kiyu, A., Hardin, S., 1993. Latrine use in rural Sarawak, Malaysia. Southeast Asian J. Trop. Med. Public Health 24, 40–42.

Kjellsson, G., Clarke, P., Gerdtham, U.G., 2014. Forgetting to remember or remembering to forget: A study of the recall period length in health care survey questions. J. Health Econ. 35, 34–46. https://doi.org/10.1016/j.jhealeco.2014.01.007

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., Nala, R., Cumming, O., Brown, J., 2020. Effects of an urban sanitation intervention on childhood enteric infection and diarrhoea in Mozambique. medRxiv Prepr. https://doi.org/https://doi.org/10.1101/2020.08.20.20178608

Knee, J., Sumner, T., Adriano, Z., Berendes, D., de Bruijn, E., Schmidt, W.P., Nalá, R., Cumming, O., Brown, J., 2018. Risk factors for childhood enteric infection in urban Maputo, Mozambique: A cross-sectional study. PLoS Negl. Trop. Dis. 12, e0006956.

https://doi.org/10.1371/journal.pntd.0006956
Koestler, L., Koestler, A., Koestler, M., Koestler, V., 2010. Improving sustainability using incentives for operation and maintenance: The concept of water-person-years. Waterlines 29, 147–162. https://doi.org/10.3362/1756-3488.2010.014
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." Qual. Life Res. 22, 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. J. Chiropr. Med. 15, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012
Kulkarni, S., O'Reilly, K., Bhat, S., 2017. No relief: lived experiences of inadequate sanitation access of poor urban women in India. Gend. Dev. 25, 167–183. https://doi.org/10.1080/13552074.2017.1331531
Kwiringira, J., Atekyereza, P., Niwagaba, C., Günther, I., 2014. Gender variations in access, choice to use and cleaning of shared latrines; experiences from Kampala Slums, Uganda. BMC Public Health 14, 1–11. https://doi.org/10.1186/1471-2458-14-1180
Lagerkvist, C.J., Kokko, S., Karanja, N., 2014. Health in perspective : Framing motivational factors for personal sanitation in urban slums in Nairobi , Kenya , using anchored best-worst scaling. https://doi.org/10.2166/wasdev.2013.069
Larsen, B., 2018a. Benefits and Costs of Drinking Water and Sanitation Interventions in Andhra Pradesh. Copenhagen Consensus - Andrha Pradesh Priorities. https://doi.org/10.1017/CBO9781139600484.012
Larsen, B., 2018b. Benefits and Costs of Drinking Water and Sanitation Interventions in Rajasthan. Copenhagen Consensus - Rajasthan Priorities.
Larsen, B., 2016. Benefits and Costs of Drinking Water, Sanitation and Hygiene Interventions. Copenhagen Consensus - Bangladesh Priorities.
Larsen, B., 2003. Hygiene and health in developing countries: Defining priorities through cost-benefit assessments. Int. J. Environ. Health Res. 13. https://doi.org/10.1080/0960312031000122172
Loomes, G., Sugden, R., 1982. Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. Econ. J. 92, 805. https://doi.org/10.2307/2232669
Lorgelly, P., Lorimer, K., Fenwick, E.A.L., Briggs, A.H., Anand, P., 2015. Operationalising the capability approach as an outcome measure in public health: The development of the OCAP-18. Soc. Sci. Med. 142, 68–81. https://doi.org/10.1016/j.socscimed.2015.08.002
Lorgelly, P.K., Lorimer, K., Fenwick, E.A.L., Briggs, A.H., Anand, P., 2015. Operationalising the capability approach as an outcome measure in public health: The development of the OCAP-18. Soc. Sci. Med. 142, 68–81. https://doi.org/10.1016/j.socscimed.2015.08.002
Lu, X., White, H., 2014. Robustness checks and robustness tests in applied economics. J. Econom. 178, 194–206. https://doi.org/https://doi.org/10.1016/j.jeconom.2013.08.016
Luby, S.P., 2017. Clean water, clean hands or new vaccines? J. Infect. 74, S18–S22. https://doi.org/10.1016/S0163-4453(17)30186-X
Luby, S.P., Rahman, M., Arnold, B.F., Unicomb, L., Ashraf, S., Winch, P.J., Stewart, C.P., Begum, F., Hussain, F., Benjamin-Chung, J., Leontsini, E., Naser, A.M., Parvez, S.M., Hubbard, A.E., Lin, A., Nizame, F.A., Jannat, K., Ercumen, A., Ram, P.K., Das, K.K., Abedin, J., Clasen, T.F., Dewey, K.G., Fernald, L.C., Null, C., Ahmed, T., Colford, J.M.,

2018. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: A cluster randomised controlled trial. Lancet Glob. Heal. https://doi.org/10.1016/S2214-109X(17)30490-4
Macuane, J.J., Buur, L., Monjane, C.M., 2018. Power, conflict and natural resources: The Mozambican Crisis revisited. Afr. Aff. (Lond). 117, 415–438. https://doi.org/10.1093/afraf/adx029
Mangham-Jefferies, L., Pitt, C., Cousens, S., Mills, A., Schellenberg, J., 2014. Cost-effectiveness of strategies to improve the utilization and provision of maternal and newborn health care in low-income and lower-middle-income countries: A systematic review. BMC Pregnancy Childbirth 14, 243. https://doi.org/10.1186/1471-2393-14-243
Maslow, A.H., 1943. A theory of human motivation. Psychol. Rev. 370–396.
Mattson, K., 2016. Final Evaluation of the Water & Sanitation for the Urban Poor JSDF Funded Maputo Peri-Urban Sanitation Project.
Mauskopf, J.A., Paul, J.E., Grant, D.M., Stergachis, A., 1998. The role of cost-consequence analysis in healthcare decision-making. Pharmacoeconomics 13, 277–288. https://doi.org/10.2165/00019053-199813030-00002
McManus, E., Turner, D., Sach, T., 2019. Can You Repeat That? Exploring the Definition of a Successful Model Replication in Health Economics. Pharmacoeconomics 37, 1371–1381. https://doi.org/10.1007/s40273-019-00836-y
Meddings, D.R., Ronald, L.A., Marion, S., Pinera, J.F., Oppliger, A., 2004. Cost effectiveness of a latrine revision programme in Kabul, Afghanistan. Bull. World Health Organ. 82, 281–289.
Meeks, R.C., 2017. Water works: The economic impact of water infrastructure. J. Hum. Resour. 52, 1119–1153. https://doi.org/10.3368/jhr.52.4.0915-7408R1
Messmer, U., 2011. Analyzing drivers and barriers to ecological sanitation by applying a sustainability framework including capabilities. University of Basel - MSc thesis.
Microsoft, 2019. Microsoft Excel for Mac.
MISAU, INE, ICF International, 2013. Moçambique Inquérito Demográfico e de Saúde 2011. Calverton, Maryland, USA.
Mitchell, P.M., Roberts, T.E., Barton, P.M., Coast, J., 2015. Assessing sufficient capability: A new approach to economic evaluation. Soc. Sci. Med. 139, 71–79. https://doi.org/10.1016/j.socscimed.2015.06.037
Mitchell, P.M., Venkatapuram, S., Richardson, J., Iezzi, A., Coast, J., 2017. Are Quality-Adjusted Life Years a Good Proxy Measure of Individual Capabilities? Pharmacoeconomics 35, 637–646. https://doi.org/10.1007/s40273-017-0495-3
Mitlin, D., Satterthwaite, D., 2013. Urban Poverty in the Global South: Scale and Nature. Routledge, London.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Group, T.P., 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLOS Med. 6, e1000097.
Mooney, G., 1989. QALYs: Are they enough? A health economist's perspective. J. Med. Ethics 15, 148–152. https://doi.org/10.1136/jme.15.3.148
Mukherjee, N., 2001. Achieving Sustained Sanitation for the Poor: Policy and Strategy Lessons from Participatory Assessments in Cambodia, Indonesia and Vietnam. World Bank.
Murray, C.J.L., Evans, D.B., Acharya, A., Baltussen, R.M.P.M., 2000. Development of WHO guidelines on generalized cost-effectiveness analysis. Health Econ. 9, 235–251.

https://doi.org/10.1002/(SICI)1099-1050(200004)9:3<235::AID-HEC502>3.0.CO;2-O
Murray, C.J.L., Lopez, A.D., 1996. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 349, 1436–1442. https://doi.org/10.1016/S0140-6736(96)07495-8
Murray, J., McFarland, D.A., Waldman, R.J., 1998. Cost-effectiveness of oral cholera vaccine in a stable refugee population at risk for epidemic cholera and in a population with endemic cholera. Bull. World Health Organ. 76, 343–352.
mWater, 2019. mWater Surveyor.
Netten, A., Burge, P., Malley, J., Potoglou, D., Towers, A.-M., Brazier, J., Flynn, T., Forder, J., Wall, B., 2012. Outcomes of social care for adults: developing a preference-weighted measure. Health Technol. Assess. (Rockv). 16. https://doi.org/10.3310/hta16160
Neumann, P.J., Thorat, T., Zhong, Y., Anderson, J., Farquhar, M., Salem, M., Sandberg, E., Saret, C.J., Wilkinson, C., Cohen, J.T., 2016. A Systematic Review of Cost-Effectiveness Studies Reporting Cost-per-DALY Averted 1–14. https://doi.org/10.1371/journal.pone.0168512
Nguyen, V.A., Hutton, G., Lan, H.T., Dan, P.H., Hoa, L.T., Nhung, B.T., 2012. Economic assessment of sanitation interventions in Vietnam. World Bank.
NICE, 2020. Developing NICE guidelines: the manual, Process and methods guides.
NICE, 2018. Guide to the processes of technology appraisal [WWW Document]. URL https://www.nice.org.uk/process/pmg19
Novotný, J., Hasman, J., Lepič, M., 2018. Contextual factors and motivations affecting rural community sanitation in low- and middle-income countries: A systematic review. Int. J. Hyg. Environ. Health 221, 121–133. https://doi.org/10.1016/j.ijheh.2017.10.018
Null, C., Stewart, C.P., Pickering, A.J., Dentz, H.N., Arnold, B.F., Arnold, C.D., Benjamin-Chung, J., Clasen, T., Dewey, K.G., Fernald, L.C.H., Hubbard, A.E., Kariger, P., Lin, A., Luby, S.P., Mertens, A., Njenga, S.M., Nyambane, G., Ram, P.K., Colford, J.M., 2018. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: A cluster-randomised controlled trial. Lancet Glob. Heal. https://doi.org/10.1016/S2214-109X(18)30005-6
Nunnally, J.C., 1978. Psychometric theory, 2nd ed. McGraw Hill, New York.
Nussbaum, M., 2011. Creating Capabilities: The Human Development Approach. Harvard University Press.
Nussbaum, M., 2001. Symposium on Amartya Sen’s philosophy: 5 Adaptive preferences and women’s options. Econ. Philos. 17, 67–88. https://doi.org/DOI: 10.1017/S0266267101000153
O’Donnell, J.C., Pham, S. V., Pashos, C.L., Miller, D.W., Smith, M.D., 2009. Health technology assessment: Lessons learned from around the world - An overview. Value Heal. 12, S1–S5. https://doi.org/10.1111/j.1524-4733.2009.00550.x
O’Mahony, J.F., Newall, A.T., van Rosmalen, J., 2015. Dealing with Time in Health Economic Evaluation: Methodological Issues and Recommendations for Practice. Pharmacoeconomics 33, 1255–1268. https://doi.org/10.1007/s40273-015-0309-4
O’Reilly, K., 2016. From toilet insecurity to toilet security: creating safe sanitation for women and girls. Wiley Interdiscip. Rev. Water 3, 19–24. https://doi.org/10.1002/wat2.1122
Ochalek, J., Lomas, J., Claxton, K., 2018. Estimating health opportunity costs in low- and middle-income countries: a novel approach and evidence from cross-country data. BMJ Glob. Heal. https://doi.org/10.1136/bmjgh-2018-000964

Oliveira, D.M. de, 2014. Saneamento básico e desenvolvimento humano: um estudo de caso no município de imperatriz/ma a partir da abordagem das capacitações. Centro Universitário UNIVATES - MSc thesis.
Parkin, D., Devlin, N., 2006. Is there a case for using visual analogue scale valuations in cost-utility analysis? Health Econ. 15, 653–664. https://doi.org/10.1002/hec.1086
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).
Peletz, R., Cock-Esteb, A., Ysenburg, D., Haji, S., Khush, R., Dupas, P., 2017. Supply and Demand for Improved Sanitation: Results from Randomized Pricing Experiments in Rural Tanzania. Environ. Sci. Technol. 51, 7138–7147. https://doi.org/10.1021/acs.est.6b03846
Peletz, R., Kisiangani, J., Ronoh, P., Cock-Esteb, A., Chase, C., Khush, R., Luoto, J., 2019. Assessing the demand for plastic latrine slabs in rural Kenya. Am. J. Trop. Med. Hyg. 101, 555–565. https://doi.org/10.4269/ajtmh.18-0888
Pigou, A.C., 1920. The Economics of Welfare.
Pitt, C., Goodman, C., Hanson, K., 2016. Economic Evaluation in Global Perspective: A Bibliometric Analysis of the Recent Literature. Health Econ. 25, 9–28. https://doi.org/10.1002/hec.3305
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. Int. J. Hyg. Environ. Health 222, 765–777. https://doi.org/https://doi.org/10.1016/j.ijheh.2019.05.004
QSR International, 2018. nVivo 12.
Radin, M., Jeuland, M., Wang, H., Whittington, D., 2020. Benefit–Cost Analysis of Community-Led Total Sanitation: Incorporating Results from Recent Evaluations. J. Benefit-Cost Anal. 1–38. https://doi.org/10.1017/bca.2020.6
Rajbhandary, J., Zwarteveen, M., Guerrero, T.A., 2019. Exploring the meaning “Sanitation Well-being” in Tha khlong, Thailand, in: The 2nd International Electronic Conference on Environmental Health Sciences. pp. 1–9.
Reddy, V.R., Kullappa, M., Rao, D.M., 2008. Social Cost-Benefit Analysis of Improved Water Quality in Rural Areas: An Exploratory Study in Coastal AP. J. Soc. Econ. Dev. 10, 68–97. https://doi.org/http://www.isec.ac.in/journal.htm
Reeser, R., 1988. Economics of Water Point Development In Central Tunisia (No. 36). Binghamton, NY.
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., Brundage, M., Ahmed, S., Aaronson, N.K., Butt, Z., 2013. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual. Life Res. 22, 1889–1905. https://doi.org/10.1007/s11136-012-0344-y
Remme, M., Siapka, M., Vassall, A., Heise, L., Jacobi, J., Ahumada, C., Gay, J., Watts, C., 2014. Review article The cost and cost-effectiveness of gender-responsive interventions for HIV : a systematic review. https://doi.org/10.7448/IAS.17.1.19228
Ritchie, J., Spencer, L., 1994. Chapter 9: Qualitative data analysis for applied policy research,

in: Burgess, R.G., Bryman, A. (Eds.), Analyzing Qualitative Data. Routledge, London ; New York, pp. 173–194.
Robeyns, I., 2017. Wellbeing, Freedom and Social Justice. https://doi.org/http://dx.doi.org/10.11647/OBP.0130
Robeyns, I., 2005. The Capability Approach: a theoretical survey. J. Hum. Dev. 6 NV-6, 93 SE-93–114. https://doi.org/10.1080/146498805200034266
Robinson, L.A., Hammitt, J.K., Cecchini, M., Chalkidou, K., Cropper, M., Hoang, P., Eozenou, V., Ferranti, D. De, Anil, B., Guanais, F., Jamison, D.T., Kwon, S., Lauer, J.A., Keeffe, L.O., Walker, D., Whittington, D., Wilkinson, T., 2019a. Reference Case Guidelines for Benefit-Cost Analysis in Global Health and Development.
Robinson, L.A., Hammitt, J.K., O ’Keeffe, L., 2019b. Valuing Mortality Risk Reductions in Global Benefit-Cost Analysis. J. Benefit Cost Anal 1–36. https://doi.org/10.1017/bca.2018.26
Rodriguez, U.-P., Hutton, G., Boatman, A., 2013. Economic assessment of sanitation interventions in Lao PDR.
Rodriguez, U.-P., Hutton, G., Jamora, N., Harder, D., Ockelford, J., Galing, E., 2011. Economic assessment of sanitation interventions in the Philippines.
Rogers, E., Tappis, H., Doocy, S., Martinez, K., Villeminot, N., Suk, A., Kumar, D., Pietzsch, S., Puett, C., 2019. Costs and cost-effectiveness of three point-of-use water treatment technologies added to community-based treatment of severe acute malnutrition in Sindh Province, Pakistan. Glob. Health Action 12, 1568827.
Ryen, L., Svensson, M., 2015. The Willingness to Pay for a Quality Adjusted Life Year: A Review of the Empirical Literature. Health Econ. 24, 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. Soc. Sci. Med. 139, 80–89. https://doi.org/10.1016/j.socscimed.2015.06.031
Saldaña, J., 2015. The Coding Manual for Qualitative Researchers, third edit. ed. Sage Publishing.
Sampson, C., 2017. Chris Sampson’s journal round-up for 9th October 2017 [WWW Document]. Acad. Heal. Econ. Blog. URL https://aheblog.com/2017/10/09/chris-sampsons-journal-round-up-for-9th-october-2017/ (accessed 9.6.20).
Sampson, C., Devlin, N., Parkin, D., 2020. Drop Dead: Is Anchoring at ‘Dead’ a Theoretical Requirement in Health State Valuation?, OHE Research Paper. London.
Samuelson, P., 1954. The Pure Theory of Public Expenditure. Rev. Econ. Stat. 36, 387–389. https://doi.org/10.2307/1925895
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. J. Water, Sanit. Hyg. Dev. 1–10. 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. Soc. Sci. Med. 217, 121–134. https://doi.org/https://doi.org/10.1016/j.socscimed.2018.09.016

Scott, R.E., Ross, I., Hawkins, P., Blackett, I., Smith, M.D., 2019. Diagnostics for assessing city-wide sanitation services. J. Water, Sanit. Hyg. Dev. 9, 111–118. https://doi.org/10.2166/washdev.2018.113
Sen, A., 2005. Human Rights and Capabilities. J. Hum. Dev. 6, 151–166. https://doi.org/10.1080/14649880500120491
Sen, A., 1999. Development as freedom. Oxford University Press.
Sen, A., 1993. Capability and Well-Being, in: Sen, Nussbaum (Eds.), The Quality of Life. Clarendon press, pp. 30–54.
Sen, A., 1985. Commodities and Capabilities. Elsevier Science Publishers, Oxford.
Sen, A., 1980. Equality of what?, in: The Tanner Lecture on Human Values. Cambridge University Press, Cambridge, pp. 197–220. https://doi.org/10.1093/0198289286.003.0002
Shemilt, I., Aluko, P., Graybill, E., Craig, D., Henderson, C., Drummond, M., Wilson, E.C.F., Robalino, S., Vale, L., Group, on behalf of the C. and C.E.M., 2019. Economic evidence. Cochrane Handb. Syst. Rev. Interv., Wiley Online Books. https://doi.org/doi:10.1002/9781119536604.ch20
Shiras, T., Cumming, O., Brown, J., Muneme, B., Nala, R., Dreibelbis, R., 2018. Shared latrines in Maputo, Mozambique: Exploring emotional well-being and psychosocial stress. BMC Int. Health Hum. Rights 18, 1–12. https://doi.org/10.1186/s12914-018-0169-z
Shrestha, R.K., Marseille, E., Kahn, J.G., Lule, J.R., Pitter, C., Blandford, J.M., Bunnell, R., Coutinho, A., Kizito, F., Quick, R., Mermin, J., R.K., S., E., M., J.G., K., J.R., L., C., P., J.M., B., R., B., A., C., F., K., R., Q., 2006. Cost-effectiveness of home-based chlorination and safe water storage in reducing diarrhea among HIV-affected households in rural Uganda. Am. J. Trop. Med. Hyg. 74, 884–890. https://doi.org/10.4269/ajtmh.2006.74.884
Simon, J., Anand, P., Gray, A., Rugkåsa, J., Yeeles, K., Burns, T., 2013. Operationalising the capability approach for outcome measurement in mental health research. Soc. Sci. Med. 98, 187–196. https://doi.org/10.1016/j.socscimed.2013.09.019
Sklar, R., 2017. Pit Latrines or Container Based Toilets?: A Cost-Benefit Analysis comparing two approaches to improving sanitation access in urban areas of Haiti.
Smith, R.D., 1996. Is Regret Theory an alternative basis for estimating the value of healthcare interventions? Health Policy (New. York). 37, 105–115. https://doi.org/10.1016/0168-8510(96)00824-X
Solomons, N., 1978. Resume of the discussion on "water and other environmental interventions." Am. J. Clin. Nutr. 31, 2124–2126.
Spears, D., 2013. Policy Lessons from the Implementation of Indias Total Sanitation Campaign. India Policy Forum 9, 63–104.
Spencer, L., Ritchie, J., O'Connor, W., 2003. Analysis: Practices, Principles and Processes, in: Ritchie, J., Lewis, J. (Eds.), Qualitative Research Practice: A Guide for Social Science Students and Researchers. Sage.
StataCorp, 2019. Stata Statistical Software: Release 16.
Stevenson, E.G.J., Greene, L.E., Maes, K.C., Ambelu, A., Tesfaye, Y.A., Rheingans, R., Hadley, C., 2012. Water insecurity in 3 dimensions: An anthropological perspective on water and women's psychosocial distress in Ethiopia. Soc. Sci. Med. 75, 392–400. https://doi.org/10.1016/j.socscimed.2012.03.022
Stiglitz, J., Sen, A., Fitoussi, J.-P., 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress.

Stillwell, W.G., Seaver, D.A., Edwards, W., 1981. A comparison of weight approximation techniques in multiattribute utility decision making. Organ. Behav. Hum. Perform. 28, 62–77. https://doi.org/https://doi.org/10.1016/0030-5073(81)90015-5

Strande, L., Schoebitz, L., Bischoff, F., Ddiba, D., Okello, F., Englund, M., Ward, B.J., Niwagaba, C.B., 2018. Methods to reliably estimate faecal sludge quantities and qualities for the design of treatment technologies and management solutions. J. Environ. Manage. 223, 898–907. https://doi.org/10.1016/j.jenvman.2018.06.100

Streeten, P., 1982. First things first : meeting basic human needs in the developing countries. Washington, D.C.

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.

Sutton, E.J., Coast, J., 2014. Development of a supportive care measure for economic evaluation of end-of-life care using qualitative methods. Palliat. Med. 28, 151–157. https://doi.org/10.1177/0269216313489368

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. J. Clin. Epidemiol. 60, 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Teschl, M., Comim, F., 2005. Adaptive Preferences and Capabilities: Some Preliminary Conceptual Explorations. Rev. Soc. Econ. 63, 229–247. https://doi.org/10.1080/00346760500130374

Tidwell, J.B., Chipungu, J., Chilengi, R., Aunger, R., 2018. Assessing peri-urban sanitation quality using a theoretically derived composite measure in Lusaka, Zambia. J. Water, Sanit. Hyg. Dev. 1–11. https://doi.org/10.2166/washdev.2018.029

Tidwell, J.B., Terris-Prestholt, F., Quaife, M., Aunger, R., 2019. Understanding demand for higher quality sanitation in peri-urban Lusaka, Zambia through stated and revealed preference analysis. Soc. Sci. Med. 232, 139–147. https://doi.org/10.1016/j.socscimed.2019.04.046

Tilley, E., Bieri, S., Kohler, P., 2013. Sanitation in developing countries: a review through a gender lens. J. Water, Sanit. Hyg. Dev. 3, 298–314. https://doi.org/10.2166/washdev.2013.090

Tong, A., Sainsbury, P., Craig, J., 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32- item checklist for interviews and focus group. Int. J. Qual. Heal. Care 19, 349–357. https://doi.org/10.1093/intqhc/mzm042

Topp, C.W., Østergaard, S.D., Søndergaard, S., Bech, P., 2015. The WHO-5 Well-Being Index: A Systematic Review of the Literature. Psychother. Psychosom. 84, 167–176. https://doi.org/10.1159/000376585

Torrance, G.W., 1986. Measurement of health state utilities for economic appraisal: A review. J. Health Econ. 5, 1–30. https://doi.org/https://doi.org/10.1016/0167-6296(86)90020-2

Torrance, G.W., Feeny, D., Furlong, W., 2001. Visual analog scales: Do they have a role in the measurement of preferences for health states? Med. Decis. Mak. https://doi.org/10.1177/02729890122062622

Turner, H.C., Lauer, J.A., Tran, B.X., Teerawattananon, Y., Jit, M., 2019a. Adjusting for Inflation and Currency Changes Within Health Economic Studies. Value Heal. 22, 1026–1032. https://doi.org/10.1016/j.jval.2019.03.021

Turner, H.C., Walker, M., Pion, S.D.S., McFarland, D.A., Bundy, D.A.P., Basáñez, M.-G.,

2019b. Economic evaluations of onchocerciasis interventions: a systematic review and research needs. Trop. Med. Int. Heal. 24, 788–816. https://doi.org/10.1111/tmi.13241
UN-HABITAT, 2010. Mozambique Cities Profile.
UNDP, 2019. Human Development Report 2019: beyond income, beyond averages, beyond today, United Nations Development Program.
UNICEF & WHO, 2019. Progress on household drinking water, sanitation and hygiene 2000-2017. Special focus on inequalities. New York.
United Nations, 2019. World Urbanization Prospects: The 2018 Revision. https://doi.org/10.4054/demres.2005.12.9
United Nations, 2017. World Population Prospects - The 2017 Revision. https://doi.org/10.1017/CBO9781107415324.004
United Nations, 2016a. The World's Cities in 2016: Data Booklet. Econ. Soc. Aff. 29. https://doi.org/10.18356/8519891f-en
United Nations, 2016b. Resolution adopted by the General Assembly on 17 December 2015: 70/169. The human rights to safe drinking water and sanitation.
van Gelder, M., 2011. Implementing Ecological Sanitation – a study on capabilities and capacities in India. University of Utrecht - MSc thesis.
van Mastrigt, G.A.P.G., Hiligsmann, M., Arts, J.J.C., Broos, P.H., Kleijnen, J., Evers, S.M.A.A., Majoie, M.H.J.M., 2016. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: a five-step approach (part 1/3). Expert Rev. Pharmacoecon. Outcomes Res. 16, 689–704. https://doi.org/10.1080/14737167.2016.1246960
Varley, R.C.G., Tarvid, J., Chao, D.N.W., 1998. A reassessment of the cost-effectiveness of water and sanitation interventions in programmes for controlling childhood diarrhoea. Bull. World Health Organ. 76, 617–631.
Vassall, A., Sweeney, S., Kahn, J.G., Gomez, G., Bollinger, L., Marseille, E., Herzel, B., Plosky, W.D., Cunnama, L., Sinanovic, E., Bautista, S., Group, G.T.A., Group, G.S., Harris, K., Levin, C., 2017. Reference Case for Estimating the Costs of Global Health Services and Interventions.
Venkatapuram, S., 2011. Health Justice: An Argument from the Capabilities Approach. Polity.
Vicary, S., 1997. Joint production and the private provision of public goods. J. Public Econ. 63, 429–445. https://doi.org/10.1016/S0047-2727(96)01602-7
von Neumann, J., Morgenstern, O., 1944. Theory of Games and Economic Behaviors. Princeton University Press.
Vyas, S., Kumaranayake, L., 2006. Constructing socio-economic status indices: How to use principal components analysis. Health Policy Plan. 21, 459–468. https://doi.org/10.1093/heapol/czl029
WageIndicator, 2016. Minimum Wages in Mozambique with effect from 01-04-2015 to 31-03-2016 [WWW Document]. URL https://wageindicator.org/salary/minimum-wage/mozambique/archive/minimum-wages-in-mozambique-with-effect-from-01-04-2015-to-31-03-2016
Wagner, E.G., Lanoix, J.N., 1958. Excreta disposal for rural areas and small communities. Monogr. Ser. World Health Organ.
Walsh, J.A., Warren, K.S., 1979. Selective Primary Health Care. N. Engl. J. Med. 301, 967–974. https://doi.org/10.1056/NEJM197911013011804
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.
Weatherly, H., Drummond, M., Claxton, K., Cookson, R., Ferguson, B., Godfrey, C., Rice, N., Sculpher, M., Sowden, A., 2009. Methods for assessing the cost-effectiveness of public health interventions: Key challenges and recommendations. Health Policy (New. York). 93, 85–92. https://doi.org/10.1016/j.healthpol.2009.07.012
Weller, S.C., Romney, A.K., 1988. Systematic Data Collection. Sage Qualitative Research Methods Series #10.
Whittington, D., Cook, J., 2018. Valuing Changes in Time Use in Low- and Middle-Income Countries. J. Benefit-Cost Anal. 10, 1–22. https://doi.org/10.1017/bca.2018.21
Whittington, D., Hanemann, W.M., Sadoff, C., Jeuland, M., 2009. The Challenge of Improving Water and Sanitation Services in Less Developed Countries. Found. Trends® Microeconomics 4, 469–609. https://doi.org/10.1561/0700000030
Whittington, D., Jeuland, M., Barker, K., Yuen, Y., 2012. Setting Priorities, Targeting Subsidies among Water, Sanitation, and Preventive Health Interventions in Developing Countries. World Dev. 40, 1546–1568. https://doi.org/10.1016/j.worlddev.2012.03.004
Whittington, D., Radin, M., Jeuland, M., 2017. Economic Costs and Benefits of three Water and Sanitation Interventions in Rural Haiti, Haïti Priorise.
WHO, 2018. Guidelines on sanitation and health, World Health Organization.
WHO, 2017. UN-Water Global Analysis and Assessment of Sanitation and Drinking-Water (GLAAS) 2017. https://doi.org/CC BY-NC-SA 3.0 IGO.
WHO, 1948. Constitution of The World Health Organization. https://doi.org/12571729
Wilbur, J., Gosling, L., 2018. Breaking the barriers – disability, ageing and HIV in inclusive WASH programming, in: Cumming, O., Slaymaker, T. (Eds.), Equality in Water and Sanitation Services. Routledge.
Wilkinson, T., Sculpher, M., Claxton, K., Revill, P., Briggs, A., Cairns, J.A., Teerawattananon, Y., Asfaw, E., Lopert, R., Culyer, A.J., Walker, D.G., 2016. The International Decision Support Initiative Reference Case for Economic Evaluation: An Aid to Thought. Value Heal. 19, 921–928. https://doi.org/10.1016/j.jval.2016.04.015
Winara, A., Hutton, G., Oktarinda, Purnomo, E., Hadiwardoyo, K., Merdykasari, I., Nurmadi, T., Bruinsma, B., Gunawan, D., Fadilah, D., Albrecht, M., 2011. Economic assessment of sanitation interventions in Indonesia.
Wolf, J., Hunter, P.R., Freeman, M.C., Cumming, O., Clasen, T., Bartram, J., Higgins, J.P.T., Johnston, R., Medlicott, K., Boisson, S., Prüss-Ustün, A., 2018. Impact of drinking water, sanitation and handwashing with soap on childhood diarrhoeal disease: updated meta-analysis and meta-regression. Trop. Med. Int. Heal. 23, 508–525. https://doi.org/10.1111/tmi.13051
World Bank, 2020. World Development Indicators [WWW Document].
World Bank, 2019a. Mozambique Urban Sanitation Project: Project Appraisal Document. https://doi.org/PIDISDSA21985
World Bank, 2019b. Health, Safety and Dignity of Sanitation Workers: An Initial Assessment. https://doi.org/10.1596/32640
World Bank, 2019c. Doing Business in Mozambique. Washington, DC.
World Bank, 2019d. 2019 Review of Staff Compensation for the World Bank Group.
World Bank, 2018. World Development Indicators [WWW Document]. URL https://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=EG (accessed 2.6.18).
World Bank, 2017. Maputo Urban Poverty and Inclusive Growth.

World Bank, 2012. Mozambique loses MZN 4 billion annually due to poor sanitation.
World Bank, 2010. Cost-Benefit Analysis of World Bank Projects.
WSUP, 2020. Referee! Responsibilities, regulations and regulating for urban sanitation.
WSUP, 2019. Systems Reboot Sanitation sector change in Maputo and Lusaka.
Wutich, A., Ragsdale, K., 2008. Water insecurity and emotional distress: Coping with supply, access, and seasonal variability of water in a Bolivian squatter settlement. Soc. Sci. Med. 67, 2116–2125. https://doi.org/https://doi.org/10.1016/j.socscimed.2008.09.042
Wyatt, A., 1990. A re-examination of costs and benefits of rural water supply projects in Central Tunisia (No. 298), WASH field report.
Yaya, S., Uthman, O.A., Amouzou, A., Bishwajit, G., 2018. Mass media exposure and its impact on malaria prevention behaviour among adult women in sub-Saharan Africa: results from malaria indicator surveys. Glob. Heal. Res. Policy 3, 20. https://doi.org/10.1186/s41256-018-0075-x
Young, S.L., Boateng, G.O., Jamaluddine, Z., Miller, J.D., Frongillo, E.A., Neilands, T.B., Collins, S.M., Wutich, A., Jepson, W., Stoler, J., 2019. The Household Water InSecurity Experiences (HWISE) Scale: development and validation of a household water insecurity measure for low-income and middle-income countries. BMJ Glob. Heal. 4, e001750. https://doi.org/10.1136/bmjgh-2019-001750
Zhang, Y., 2006. Urban-Rural Literacy Gaps in Sub-Saharan Africa: The Roles of Socioeconomic Status and School Quality. Comp. Educ. Rev. 50, 581–602. https://doi.org/10.1086/507056

Appendix A: Supplementary materials for Chapter 3 (systematic review)

A. Results tables

B. List of low- and middle-income countries in 2018

C. Search strategy

D. Excluded studies

E. Scoring protocol for CHEERS

F. Original and inflation-adjusted results of cost-effectiveness studies

G. Uncertainty and heterogeneity

H. Types of benefit and valuation methods

A. Results tables

Table 2: Characteristics of included studies - empirical interventions which were actually implemented

* "Primary" denotes data collected about the intervention by the researchers, or from expenditure records of that same intervention. For outcomes, semi-primary data is from the setting, but may have been collected in relation to a similar intervention and/or by others. For costs, semi-primary denotes expenditure records of programmes in general, or costs that were budgeted rather than spent. CBA = cost-benefit analysis; CEA = cost-effectiveness analysis; CLTS = community-led total sanitation; HWT = household water treatment; n/a = not applicable; OD = open defecation; RCT = randomised controlled trial.

Reference	ID	Country & setting	Population (if primary outcome collected)	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Study design: cost-benefit analysis								
Eklund (1991)	1184	Ethiopia (small town & rural)	n/a	Water supply – spring-fed gravity scheme piped to tapstands	Empirical - fully-costed infrastructure delivery	Societal	CBA	Costs primary, outcomes semi-primary (estimate from pre-intervention observational data).
Abelin (1997)	1181	India (rural)	n/a	Water supply – handpumps (on boreholes and converted stepwells)	Empirical - partially-costed infrastructure delivery	Societal	CBA	Both semi-primary (cross-sectional data of similar intervention). Excludes programme management costs.
Rodriguez (2011)	1069	Philippines (Urban & rural)	18,452 adults across 3 rural sites and 3 urban	Sanitation – (i) pit latrine, (ii) flush to pit, (iii) flush to septic tank (iv) urine-diverting dry toilet, (v) sewerage	Empirical - various fully-costed programmes (varied by site)	Societal	CBA & CEA	Costs mixed (cross-sectional and records), outcomes mixed (cross-sectional and literature)
Winara (2011)	1070	Indonesia (Urban & rural)	1,500 adults across 2 rural sites and 3 urban	Sanitation – (i) private pit latrine, (ii) community toilet, (iii) shared toilet, (iv) private septic tank, (v) private communal sewerage, (vi) private sewerage + treatment	Empirical - various fully-costed programmes (varied by site)	Societal	CBA & CEA	Costs mixed (cross-sectional and records), outcomes mixed (cross-sectional and literature)

Reference	ID	Country & setting	Population (if primary outcome collected)	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Chuan (2012)	1081	China (Urban & rural)	909 adults across 3 rural sites and 5 urban	Sanitation – (i) pit latrine, (ii) flush to septic tank (iii) biogas digester toilet, (iv) sewerage	Empirical - various fully-costed programmes (varied by site)	Societal	CBA & CEA	Costs mixed (cross-sectional and records), outcomes mixed (cross-sectional and literature)
Nguyen (2012)	1082	Vietnam (Urban & rural)	1,851 adults across 9 rural sites and 8 urban	Sanitation – (i) flush to pit, (ii) flush to septic tank (iii) urine-diverting dry toilet, (iv) decentralised wastewater treatment, (v) conventional sewerage	Empirical - various fully-costed programmes (varied by site)	Societal	CBA & CEA	Costs mixed (cross-sectional and records), outcomes mixed (cross-sectional and literature)
Heng (2012)	1087	Cambodia (Urban & rural)	1,172 adults across 4 rural sites and 1 urban	Sanitation – (i) pit latrine, (ii) pour-flush pit latrines, (iii) sewerage	Empirical - various fully-costed programmes (varied by site)	Societal	CBA & CEA	Costs mixed (cross-sectional and records), outcomes mixed (cross-sectional and literature)
Rodriguez (2013)	1097	Laos (Urban & rural)	1,211 adults across 3 rural sites and 3 urban	Sanitation – (i) pit latrine, (ii) flush to pit, (iii) flush to septic tank	Empirical - various fully-costed programmes (varied by site)	Societal	CBA & CEA	Costs mixed (cross-sectional and records), outcomes mixed (cross-sectional and literature)
Dickinson (2015)	1109	India (rural)	n/a	Sanitation – unspecific, typically pour-flush to off-set pit	Empirical - partially-costed CLTS with partial hardware subsidy (for people below the poverty line)	Household (excludes provider costs)	CBA	Costs primary (RCT), outcomes secondary. Excludes programme management costs.
Meeks (2017)	1128	Kyrgyzstan (rural)	776 adults	Water supply – public taps	Empirical - infrastructure delivery with unclear costing	Societal	CBA	Costs secondary, outcomes semi-primary (regression on existing observational panel data). Unclear which costs included/excluded
Cha (2018)	1143	Ghana (rural)	600 adult caregivers	Water supply – new or rehabilitated boreholes with handpump	Empirical - fully-costed infrastructure delivery	Societal	CBA	Both primary (costs from records/survey, outcomes from RCT)

Reference	ID	Country & setting	Population (if primary outcome collected)	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Burt (2018)	1145	India (urban)	4,000 adults	Water supply – upgrading existing piped network to continuous water supply (from intermittent)	Empirical - fully-costed infrastructure delivery	Societal	CBA	Costs semi-primary (programme records & budgets), outcomes primary (matched cohort)
Hutton (2020)	1173	India (rural)	18,376 adults across 12 states	Sanitation – Flush to off-set pit (51%) and flush to septic tank (38%).	Empirical - fully-costed campaign with partial ex post reward / subsidy (Swachh Bharat Mission)	Societal	CBA	Costs mixed (cross-sectional and government records), outcomes mixed (cross-sectional and literature). Programme management costs unclear.
Cha (2020)	1185	Ethiopia (rural)	906 caregivers of u5 children	Sanitation – improved pit latrine	Empirical - fully-costed CLTS with no subsidy	Societal	CBA	Both primary (costs from records/survey, outcomes from RCT)
Study design: cost-effectiveness analysis								
Meddings (2004)	1032	Afghanistan (urban)	under-5 children: 1,238 cases and 625 controls	Sanitation – vault latrines with night soil	Empirical - fully-costed infrastructure delivery at partial hardware subsidy (full latrine installations or upgrades, at c.50% subsidy for new construction and 75% for upgrades)	Societal & provider	CEA	Both primary (costs from records/survey, outcomes from case control study)
Shrestha (2006)	1036	Uganda (rural)	196 adults in HIV-affected households	HWT – 20L storage container with chlorine HWT	Empirical - fully-costed product provision	Health provider	CEA	Both primary (costs from records/survey, outcomes from RCT)
Spears (2013)	1099	India (rural)	n/a	Sanitation – unspecific, typically pour-flush to off-set pit	Empirical - partially-costed campaign with partial hardware subsidy (Total Sanitation Campaign with *ex post* reward)	Societal & Provider	CEA	Costs semi-primary (government records), outcomes secondary. Excludes operational costs.
Bhalotra (2017)	1136	Mexico (urban)	n/a	Source water treatment – source chlorination in existing piped networks	Empirical - infrastructure delivery with unclear costing	Unclear, appears provider	CEA	Costs semi-primary (unclear records), outcomes secondary. Unclear which costs included/excluded

Reference	ID	Country & setting	Population (if primary outcome collected)	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Rogers (2019)	1158	Pakistan (rural)	901 under-5 children	HWT – All in addition to standard care for severe acute malnutrition (including hygiene and water storage container): (i) chlorine disinfection, (ii) flocculant / disinfectant, (iii) ceramic filters.	Empirical - fully-costed product provision	Health provider	CEA	Both primary (costs from records/survey, outcomes from RCT)
Dupas (2020)	1179	Malawi (rural)	2,313 adults	HWT – chlorine solution (delivered free either via coupons redeemable at shops, or via community health workers)	Empirical - partially costed product provision	Health provider	CEA	Costs secondary, outcomes primary (RCT). Excludes programme management costs.

Table 3: Characteristics of included studies – hypothetical interventions which were not actually implemented

* "Primary" denotes data collected about the intervention by the researchers, or from expenditure records of that same intervention. For outcomes, semi-primary data is from the setting, but may have been collected in relation to a similar intervention and/or by others. For costs, semi-primary denotes expenditure records of programmes in general, or costs that were budgeted rather than spent. CBA = cost-benefit analysis; CEA = cost-effectiveness analysis; CLTS = community-led total sanitation; HWT = household water treatment; n/a = not applicable; OD = open defecation; RCT = randomised controlled trial.

Reference	ID	Country & setting	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Study design: cost-benefit analysis							
Reeser (1988)	1180	Tunisia (rural)	Water supply – diesel-powered deep boreholes (~300m) access off-plot by households or purchased on-plot from vendors	Hypothetical - partially-costed infrastructure delivery	Societal	CBA	Both semi-primary (observational data of similar intervention). Cost data sources cited. Excludes programme management costs.

Reference	ID	Country & setting	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Wyatt (1990)	1014	Tunisia (rural)	Water supply – diesel-powered deep boreholes (~300m) access off-plot by households or purchased on-plot from vendors	Hypothetical - fully-costed infrastructure delivery	Societal	CBA	Both semi-primary (observational data of similar intervention). Cost data sources cited.
Hutton (2004)	1030	Generic (global regions)	Water supply, sanitation and HWT – scenarios including: (i) basic water (ii) basic water & chlorine HWT, (iii) piped water and sewerage	Hypothetical - partially-costed infrastructure with unclear delivery/subsidy	Societal	CBA	Both secondary. Cost data sources cited. Excludes programme management costs.
Ignacio (2006)	1039	Philippines (urban & rural)	Sanitation – (i) ecosan, (ii) pit latrine, (iii) pour-flush to septic tank	Hypothetical - partially costed infrastructure delivery	Societal	CBA	Both secondary. Cost data sources cited. Excludes programme management costs.
Hutton (2007b)	1045	Generic (global regions)	Water supply and sanitation – (i) basic water, (ii) basic sanitation, (iii) basic WSS	Hypothetical - partially-costed infrastructure with unclear delivery	Societal	CBA	Both secondary. Cost data sources cited. Excludes programme management costs.
Hutton (2007a)	1046	Generic (global regions)	Water supply, sanitation and HWT – (i) basic water (ii) basic WSS, (iii) basic WSS & chlorine HWT (iv) piped water and sewerage	Hypothetical - partially-costed infrastructure with unclear delivery	Societal	CBA	Both secondary. Cost data sources cited. Excludes programme management costs.
Reddy (2008)	1054	India (rural)	Source water treatment – village-level UV water treatment plants, accessed pay-per-use	Hypothetical - infrastructure delivery with unclear costing.	Societal	CBA	Both semi-primary (observational data of similar intervention). Unclear which costs included.
Whittington (2009)	1055	Generic (rural)	Water supply, sanitation and HWT – Water supply: borehole with handpump. Sanitation: improved latrine, HWT : biosand filters	Hypothetical - fully-costed. Water supply: infrastructure delivery. Water treatment: product provision. Sanitation: CLTS with no subsidy	Societal	CBA	Both secondary. Cost data sources cited.
Jeuland (2009)	1059	Generic (rural)	Water supply and HWT – Water supply: borehole with handpump. Water treatment: biosand filters	Hypothetical - fully-costed product provision/ infrastructure delivery	Societal	CBA	Both secondary. Cost data sources cited.

Reference	ID	Country & setting	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Whittington (2012)	1084	Generic (rural)	Sanitation and HWT – HWT: biosand filters. Sanitation: improved latrine	Hypothetical - fully-costed. Water treatment: product provision. Sanitation: CLTS with no subsidy	Societal	CBA	Both secondary. Cost data sources cited.
Fahimuddin (2012)	1085	India (rural)	Water supply – Public taps and private connections	Hypothetical - fully costed infrastructure delivery	Societal	CBA	Costs secondary, outcomes semi-primary (cross-sectional data of similar intervention). Cost data sources cited.
Hutton (2013)	1089	Generic (global regions)	Water supply and sanitation – Rural: borehole with handpump, flush to pit latrine. Urban: piped water, flush to septic tank / sewer	Hypothetical - partially-costed infrastructure with unclear delivery	Societal	CBA	Both secondary. Cost data sources cited. Unclear which costs included.
Cronin (2014)	1105	India (urban & rural)	Sanitation – unspecific, typically pour-flush to off-set pit	Hypothetical - campaign apparently fully-costed, with partial ex post reward / subsidy but unclear delivery.	Societal	CBA	Both secondary. Cost data sources cited.
Boije (2014)	1108	India (rural)	Water supply and HWT – (i) atmospheric water generator (ii) atmospheric heat-pump (iii) solar disinfection	Hypothetical - partially costed infrastructure delivery	Societal	CBA	Both secondary. Cost data sources cited. Excludes programme management and O&M costs
Hutton (2015)	1112	Generic (global)	Water supply and sanitation – (i) urban and rural water - mix of borehole with handpump and hand-dug well, (ii) rural sanitation - mix of pour-flush to pit and unimproved pit, (iii) urban sanitation - mix septic tank, sewerage, and pit latrine.	Hypothetical - fully-costed infrastructure with unclear delivery	Societal	CBA	Both secondary. Cost data review undertaken but sources not cited.
Larsen (2016)	1122	Bangladesh (rural)	Water supply and sanitation – (i) water supply - deep tubewell, pond sand filter, rainwater harvesting. (ii) sanitation - improved pit latrine	Hypothetical - partially-costed infrastructure with unclear delivery	Societal	CBA	Both secondary. Cost data sources cited.

Reference	ID	Country & setting	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Whittington (2017)	1131	Haiti (rural)	Water supply, sanitation and HWT – (i) borehole with handpump; (ii) borehole with handpump & biosand filter; (iii) improved pit latrine	Hypothetical - fully-costed. Water supply & HWT: product/infrastructure delivery. Sanitation: CLTS with no subsidy	Societal	CBA	Both secondary. Cost data sources cited.
Sklar (2017)	1132	Haiti (urban)	Sanitation – (i) pour-flush to septic tank (ii) container-based sanitation	Hypothetical - unclear which costs included. (i) infrastructure delivery (ii) lease programme	Societal	CBA	Both secondary. Cost data sources cited. Unclear which costs included.
Larsen (2018a)	1140	India (urban & rural)	Water supply, sanitation and HWT – (i) water supply - piped water (urban) and point sources (rural); (ii) sanitation - improved sanitation (urban and rural), campaign promoting use of existing sanitation by all household members (iii) HWT - water filters	Hypothetical - partially costed. Water supply and HWT - product/infrastructure at full subsidy. Sanitation - toilet use campaign with no subsidy.	Societal	CBA	Both secondary. Cost data sources not cited. Excludes programme management costs.
Larsen (2018b)	1149	India (urban & rural)	Water supply, sanitation and HWT – (i) water supply - piped water (urban) and point sources (rural); (ii) sanitation - improved sanitation (urban and rural), campaign promoting use of existing sanitation by all household members; (iii) HWT - water filters	Hypothetical - partially costed. Water supply and HWT – product / infrastructure delivery. Sanitation - campaign to use existing toilets	Societal	CBA	Both secondary. Cost data sources not cited. Excludes programme management costs.
Dasgupta (2019)	1162	India (rural)	Sanitation – eco-san toilet and twin-pit pour flush toilet (with latter assumed to receive subsidy under Swachh Bharat Mission national programme)	Hypothetical - partially costed infrastructure delivery	Societal	CBA	Costs primary (cross-sectional), outcomes secondary. Cost data sources cited. Excludes programme management costs.
Das (2019)	1167	Bangladesh (rural)	Water supply – pond-sand filters (to reduce salinity)	Hypothetical - partially-costed infrastructure with unclear delivery	Societal	CBA	Costs secondary, outcomes semi-primary (regression on cross-sectional data). Cost data sources cited.
Radin (2020)	1175	Generic (rural)	Sanitation – improved pit latrine	Hypothetical - fully-costed CLTS with no subsidy	Societal	CBA	Both secondary. Cost data sources cited.

Reference	ID	Country & setting	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Dwumfour-Asare (2020)	1178	Ghana (urban)	Sanitation – (i) full subsidy of biodigestor toilets with bylaw enforcement (ii) subsidy only	Hypothetical - fully-costed campaign with full subsidy and with/without enforcement	Societal	CBA	Both secondary. Most cost data sources uncited.
Study design: cost-effectiveness analysis							
Varley (1998)	1020	Generic (urban)	Water supply – public taps	Hypothetical - partially-costed infrastructure with unclear delivery	Unclear, appears provider	CEA	Both secondary. Cost data sources not cited.
Murray (1998)	1183	Generic, (i) endemic cholera, (ii) epidemic cholera	Water supply and sanitation – stand pipes and pit latrines	Hypothetical - partially-costed infrastructure with unclear delivery	Societal	CEA	Both secondary. Cost data sources cited. Programme management costs excluded.
Jha (1998)	1022	Guinea (national)	Water supply and sanitation – pit latrines and "safe water"	Hypothetical - infrastructure with unclear delivery and costing.	Unclear, appears provider	CEA	Both secondary. Cost data sources not cited. Unclear which costs included.
Larsen (2003)	1029	Generic (global regions)	Water supply and sanitation – "safe water and sanitation"	Hypothetical - partially-costed infrastructure with unclear delivery	Societal	CEA	Both secondary. Cost data sources cited. Recurrent and programme management costs excluded.
Cairncross (2006)	1037	Generic (national)	Water supply and sanitation: hand pump or standpost, network connection, improved latrine	Hypothetical - partially costed. Water supply: (i) infrastructure delivery, (ii) regulation and advocacy. Sanitation: (i) infrastructure delivery with promotion, (ii) promotion without subsidy.	Societal	CEA	Both secondary. Cost data sources cited.
Clasen (2007)	1043	Generic (global regions)	Water supply and HWT – scenarios including: (i) wells, boreholes and standpoints, (ii) chlorine HWT, (iii) filtration HWT, (iv) solar disinfection HWT, (v) flocculation/disinfection HWT	Hypothetical - partially costed product provision	Societal	CEA	Both secondary. Cost data sources cited. Unclear which costs included.

Reference	ID	Country & setting	Service/technology	Type of intervention and extent of costing (see Table 1)	Costing perspective	Study design	Data sources for costs and benefits*
Haller (2007)	1044	Generic (global regions)	Water supply, sanitation and HWT – (i) basic WSS, (ii) chlorine HWT, (iii) basic water & chlorine HWT, (iv) piped water & sewerage	Hypothetical - partially-costed infrastructure with unclear delivery	Societal	CEA	Both secondary. Cost data sources cited. Excludes programme management costs.
Barungi (2011)	1071	Uganda (urban)	Water supply – public stand-posts and boreholes	Hypothetical - partially costed infrastructure delivery	Societal	CEA	Both secondary. Cost data sources cited. Excludes programme management and O&M costs
Gunther (2013)	1092	Generic (40 countries)	Water supply and sanitation – (i) off-plot water and pit latrine, (ii) on-plot piped and flush to septic tank / sewer	Hypothetical - partially-costed infrastructure with unclear delivery	Provider	CEA	Both secondary. Cost data sources cited. Excludes programme management and O&M costs.

Table 4: Study results, empirical cost-benefit analyses

*CHEERS scores are a proxy for study quality. ** Refer to the individual publications for specific results ranges under different kinds of DSA and PSA. ACER = average cost-effectiveness ratio; DSA = deterministic sensitivity analysis; PSA = probabilistic sensitivity analysis; OD = open defecation; ICER = incremental cost-effectiveness ratio.

Reference	ID	Study quality*	Comparator	Intervention and technology	Base case results Reported as BCRs for CBA, 2019 US$ for CEA	Sensitivity analysis**
Eklund (1991)	1184	medium	Without intervention (ponds, rivers, rainwater)	Delivery of gravity schemes with tapstands	Positive NPV in some scenarios but not others	DSA
Abelin (1997)	1181	high	Without intervention (open step-wells and ponds)	Delivery of handpumps	Positive NPV	DSA
Rodriguez (2011)	1069	medium	Without intervention (OD)	Multiple intervention types, resulting in: (i) rural dry pit, (ii) rural flush to pit, (iii) rural flush to septic tank, (iv) urban flush to pit, (v) urban flush to septic tank, (vi) urban flush to sewer	(i) 4.7 (ii) 5.7 (iii) 2.3 (iv) 2.8 (v) 4.5 (vi) 1.9	DSA

Reference	ID	Study quality*	Comparator	Intervention and technology	Base case results Reported as BCRs for CBA, 2019 US$ for CEA	Sensitivity analysis**
Winara (2011)	1070	medium	Without intervention (OD / unimproved)	Multiple intervention types, resulting in: (i) rural flush to pit, (ii) rural flush to septic tank, (iii) urban flush to pit, (iv) urban flush to septic tank, (v) urban flush to sewer	(i) 5.6 (ii) 3.2 (iii) 2.4 (iv) 1.7 (v) 1.1	DSA
Chuan (2012)	1081	medium	Without intervention (OD)	Multiple intervention types, resulting in: (i) rural dry pit, (ii) rural flush to septic tank, (iii) urban dry pit, (iv) urban flush to septic tank, (v) urban flush to sewer	(i) 8.5 (ii) 4.7 (iii) 5.4 (iv) 2.8 (v) 1.9	DSA
Nguyen (2012)	1082	medium	Without intervention (OD)	Multiple intervention types, resulting in: (i) rural flush to pit, (ii) rural flush to septic tank, (iii) urban flush to pit, (iv) urban flush to septic tank, (v) urban flush to sewer	(i) 6.4 (ii) 3.4 (iii) 6.8 (iv) 2.9 (v) 2.4	DSA
Heng (2012)	1087	medium	Without intervention (OD)	Multiple intervention types, resulting in: (i) rural dry pit, (ii) rural flush to pit, (iii) urban flush to pit, (iv) urban flush to sewer	(i) 1.3 (ii) 1.8 (iii) 1.4 (iv) 0.03	DSA
Rodriguez (2013)	1097	medium	Without intervention (OD)	Multiple intervention types, resulting in: (i) rural dry pit, (ii) rural flush to pit, (iii) rural flush to septic tank, (iv) urban dry pit, (v) urban flush to pit	(i) 7.2 (ii) 6.6 (iii) 3.7 (iv) 5.9 (v) 2.2	DSA
Dickinson (2015)	1109	medium	Before intervention (primarily OD)	CLTS with subsidy - typically achieved pour-flush to off-set pit	3.0	none
Cha (2018)	1143	very high	Control group (unclear but appears unimproved water)	Delivery of (i) new boreholes with handpump, (ii) rehabilitated boreholes with handpump	(i) 9.4 (ii) 14.1	DSA
Burt (2018)	1145	very high	Without intervention (intermittent supply in other parts of the city)	Delivery of upgraded water network with continuous supply	Negative NPV for the most realistic scenarios	DSA
Hutton (2020)	1173	high	Before intervention ("recalled situation prior to owning a toilet. For the majority of HHs this was OD")	Campaign with ex post subsidy, resulting mostly in flush to offset pit / septic tank. Results for (i) all households, (ii) poorest quintile, (iii) richest quintile	(i) 3.0 (ii) 4.2 (iii) 2.4	DSA & PSA
Cha (2020)	1185	very high	Existing government approach (very limited promotion via health extension workers)	CLTS with improved pit latrine	3.7	DSA & PSA

Table 5: Study results, hypothetical cost-benefit analyses

*CHEERS scores are a proxy for study quality. ** Refer to the individual publications for specific results ranges under different kinds of DSA and PSA. ACER = average cost-effectiveness ratio; DSA = deterministic sensitivity analysis; PSA = probabilistic sensitivity analysis; OD = open defecation; ICER = incremental cost-effectiveness ratio.

Reference	ID	Study quality*	Comparator	Intervention and technology	Base case results Reported as BCRs for CBA, 2019 US$ for CEA	Sensitivity analysis**
Reeser (1988)	1180	high	Without intervention (existing use of vendors or further away water points)	Delivery of diesel-powered deep boreholes	1.0	DSA
Wyatt (1990)	1014	very high	Without intervention (vendors and/or further-off water points)	Delivery of diesel-powered deep boreholes	1.3	DSA
Hutton (2004)	1030	medium	Without intervention (existing regional coverage)	Delivery of: (i) Water supply, (ii) WSS (iii) WSS + HWT, (iv) piped water and sewerage	(i) 10.1 (ii) 10.0 (iii) 13.4 (iv) 4.0	DSA
Ignacio (2006)	1039	medium	Without intervention (unclear, appears OD)	Delivery of: (i) Basic ecosan, (ii) Concrete ecosan, (iii) Pit toilet, (iv) Pour-flush to ST	(i) 0.9 (ii) 0.5 (iii) 1.0 (iv) 0.4	none
Hutton (2007b)	1045	medium	Without intervention (regions continue existing trajectory)	Delivery of: (i) water, (ii) sanitation, (iii) WSS	(i) 5.8 (ii) 11.2 (iii) 10.3	DSA
Hutton (2007a)	1046	medium	Without intervention (existing regional coverage)	Delivery of: (i) water supply, (ii) WSS (iii) WSS + HWT, (iv) piped water and sewerage	(i) 4.4 (ii) 5.2 (iii) 5.7 (iv) 2.1	DSA
Reddy (2008)	1054	low	Without intervention (public taps)	Delivery of village-level UV water treatment plants	0.86-3.00	DSA
Whittington (2009)	1055	high	Without intervention: (i) further away "traditional" sources (ii) OD, (iii) off-plot improved or unimproved	(i) Delivery of boreholes with hand pump (ii) CLTS with basic pit latrine (iii) Provision of biosand filters	(i) 3.1 (ii) 2.8 (iii) 2.8	DSA & PSA
Jeuland (2009)	1059	high	Without intervention: (i) further away "traditional" sources (ii) off-plot improved or unimproved	Delivery of (i) boreholes with hand pump, (ii) biosand filters	(i) 3.2 (ii) 2.9	DSA & PSA
Whittington (2012)	1084	high	Without intervention: (i) OD, (ii) off-plot improved or unimproved	(i) CLTS with basic pit latrine (ii) Provision of biosand filters	(i) 2.2 (ii) 5.7	DSA & PSA

Reference	ID	Study quality*	Comparator	Intervention and technology	Base case results Reported as BCRs for CBA, 2019 US$ for CEA	Sensitivity analysis**
Fahimuddin (2012)	1085	high	Without intervention (mix of wells, springs, rivers and public taps)	Delivery of public taps and private connections	2.0	none
Hutton (2013)	1089	medium	Without intervention (JMP data per region)	Delivery of: (i) Water supply (see Table 2) (ii) Sanitation (see Table 2)	(i) 2.0 (ii) 5.5	DSA
Cronin (2014)	1105	low	Without intervention (primarily OD)	Unclear intervention resulting in pour-flush to off-set pit	5.7	DSA
Boije (2014)	1108	medium	Without intervention (unclear)	Delivery of (i) atmospheric water generator (ii) atmospheric heat-pump, (iii) solar disinfection	(i) 3.9 (ii) 0.9 (iii) 10.3	DSA
Hutton (2015)	1112	medium	Without intervention (JMP data per country, only for people using unimproved)	Delivery of (i) rural water (see Table 2), (ii) rural sanitation (see Table 2)	(i) 6.8 (ii) 5.2	DSA
Larsen (2016)	1122	medium	Without intervention (shallow tubewell)	Delivery of (i) deep tubewell, (ii) pond sand filter, (iii) rainwater harvesting, (iv) improved pit latrine (from unimproved), (v) improved pit latrine (from shared)	(i) 10.0 (ii) 7.0 (iii) 8.0 (iv) 1.4 (v) 2.1	DSA
Meeks (2017)	1128	medium	Before intervention (unclear)	Delivery of public taps	positive NPV (implied BCR of 8.2 not reported)	DSA
Whittington (2017)	1131	very high	Without intervention (unimproved water and OD)	(i) Delivery of boreholes with handpump (ii) Delivery of boreholes with handpump & biosand filter (iii) Promotion: CLTS with improved pit latrine	(i) 2.2 (ii) 2.1 (iii) 1.1	DSA & PSA
Sklar (2017)	1132	high	Without intervention (population using unimproved services, predominantly shared or low-quality toilets)	(i) Delivery of pour-flush to septic tank (ii) Lease of container-based sanitation	(i) 0.90 (ii) 0.99	DSA
Larsen (2018)	1140	medium	Without intervention (varies per intervention but not always clear)	(i) Delivery of urban improved water supply, (ii) Delivery of rural improved water supply, (iii) Delivery of urban improved sanitation, (iv) Delivery of rural improved sanitation, (v) Campaign to use existing toilets, (vi) Promotion of HWT filter	(i) 11.6 (ii) 5.7 (iii) 7.2 (iv) 9.0 (v) 2.2 (vi) 8.2	DSA

Reference	ID	Study quality*	Comparator	Intervention and technology	Base case results Reported as BCRs for CBA, 2019 US$ for CEA	Sensitivity analysis**
Larsen (2018)	1149	medium	Without intervention (varies per intervention but not always clear)	(i) Delivery of urban improved water supply, (ii) Delivery of rural improved water supply, (iii) Delivery of urban improved sanitation, (iv) Delivery of rural improved sanitation, (v) Campaign to use existing toilets, (vi) Promotion of HWT filter	(i) 8.9 (ii) 4.1 (iii) 6.2 (iv) 7.8 (v) 1.8 (vi) 6.8	DSA
Dasgupta (2019)	1162	high	Without intervention (OD)	Delivery of (i) eco-san, (ii) twin-pit pour flush toilet	(i) 3.3 (ii) 1.7	DSA
Das (2019)	1167	medium	Without intervention (tubewells)	Delivery of pond-sand filters	Positive NPV in all scenarios	DSA
Radin (2020)	1175	very high	Without intervention (assumes 45% baseline toilet coverage based on review of CLTS trials)	CLTS with improved pit latrine in (i) low-uptake villages (5 percentage point (pp) coverage increase), (ii) medium-uptake villages (15pp), (iii) high-uptake villages (35pp)	(i) 0.7 (ii) 1.7 (iii) 3.2	DSA & PSA
Dwumfour-Asare (2020)	1178	medium	Without intervention (breakdown of service use before/after is on p.12)	Campaign with delivery of (i) fully subsidised biodigestor toilets with bylaw enforcement, at 75% uptake, (ii) subsidy only, at 10% uptake	(i) 4.2 (ii) 5.2	DSA

Table 6: Study results, empirical cost-effectiveness analyses

*CHEERS scores are a proxy for study quality. ** Refer to the individual publications for specific results ranges under different kinds of DSA and PSA. ACER = average cost-effectiveness ratio; DSA = deterministic sensitivity analysis; PSA = probabilistic sensitivity analysis; OD = open defecation; ICER = incremental cost-effectiveness ratio.

Reference	ID	Study quality*	Comparator	Intervention and technology	Base case results Reported as BCRs for CBA, 2019 US$ for CEA	Sensitivity analysis**
Meddings (2004)	1032	high	Control - unimproved toilets ("the original latrine is destroyed")	Delivery of vault latrines	per death averted (ACER): 4,862 (provider perspective) 9,260 (societal perspective)	none

Reference	ID	Study quality*	Comparator	Intervention and technology	Base case results Reported as BCRs for CBA, 2019 US$ for CEA	Sensitivity analysis**
Shrestha (2006)	1036	very high	Control group - same intervention (i.e. education, antibiotics) but without HWT	Provision of 20L storage containers with chlorine	2,041 per DALY averted (ICER)	DSA
Spears (2013)	1099	medium	practice in districts without intervention (primarily OD)	Total sanitation campaign, typically achieving pour-flush to off-set pit	per infant death averted (ACER): 2,761 per life year saved (ACER): 34	DSA
Bhalotra (2017)	1136	low	practice before intervention (unchlorinated piped water)	Source chlorination of existing piped networks	per death averted (ACER): 66,175 per life year saved (ACER): 1,103	none
Rogers (2019)	1158	very high	Control - standard of care (hygiene and water container)	Alongside treatment for severe acute malnutrition, provision of: (i) chlorine disinfection, (ii) flocculant / disinfectant, (iii) ceramic filters.	Per additional child recovered (with ICERs relative to control): (i) ICER 19 , ACER 100 (ii) ICER 120, ACER 131 (iii) ICER 529, ACER 232	DSA & PSA
Dupas (2020)	1179	low	Control, but unclear	Provision of chlorine solution, delivered free via (i) coupons, (ii) community health workers (CHWs) at 40% coverage, (iii) CHWs at 100% coverage	per diarrhoea case averted (ACER) (i) $5.21 (ii) $29.02 (iii) $21.27	none

Table 7: Study results, hypothetical cost-effectiveness analyses

*CHEERS scores are a proxy for study quality. ** Refer to the individual publications for specific results ranges under different kinds of DSA and PSA. ACER = average cost-effectiveness ratio; DSA = deterministic sensitivity analysis; PSA = probabilistic sensitivity analysis; OD = open defecation; ICER = incremental cost-effectiveness ratio.

Reference	ID	Study quality*	Comparator	Intervention and technology	Base case results Reported as BCRs for CBA, 2019 US$ for CEA	Sensitivity analysis**
Varley (1998)	1020	medium	Practice without intervention ("inadequate or no hardware")	Delivery of public taps	1,770 per DALY averted (ACER) 61,017 per death averted (ACER)	DSA
Murray (1998)	1183	medium	Practice without intervention (unclear)	Delivery of (i) WASH, (ii) WASH + vaccine, (iii) WASH + patient treatment	per DALY averted (i) stable: 665, refugee: 424, (ii) stable: 854, refugee: 481 (iii) stable: 502, refugee: 261	none
Cairncross (2006)	1037	low	practice without intervention (implied OD and unimproved water)	(i) Delivery of hand pump or standpost, (ii) Delivery of house connections, (iii) Water regulation and advocacy, (iv) Delivery with campaign: pit latrine, (v) campaign only: pit latrine	per DALY averted (ACER) (i) 135 (ii) 321 (iii) 68 (iv) 389 (v) 16	none
Clasen (2007)	1043	medium	practice without intervention (existing regional coverage)	Provision of (i) wells, boreholes and standposts, (ii) Household chlorination, (iii) Household filtration, (iv) Household solar disinfection, (v) Household flocculation / disinfection	per DALY averted (ACER) (i) 171 (ii) 74 (iii) 197 (iv) 85 (v) 655	none
Haller (2007)	1044	high	practice without intervention (existing regional coverage)	Delivery of (i) HWT chlorine (on unimproved) (ii) universal WSS, (iii) universal WSS with HWT, (iv) universal piped water & sewerage	per DALY averted (ACER) (i) 29 (ii) 822 (iii) 299 (iv) 989	DSA
Barungi (2011)	1071	low	practice without intervention (unclear)	Delivery of (i) stand-posts, (ii) boreholes	per DALY averted (ACER) (i) 22,457 (ii) 35,548	DSA
Jha (1998)	1022	low	Practice without intervention (OD and unimproved water)	Delivery of pit latrines and "safe water"	467 per life year saved (ACER)	None
Larsen (2003)	1029	low	Practice without intervention (OD and unimproved water)	Delivery of (i) "Water", (ii) "Sanitation"	per child death averted (ACER): (i) 1,440 (ii) 3,072	none
Gunther (2013)	1092	medium	practice without intervention (DHS data per country)	Delivery of (i) Basic, off-plot water and pit latrine, (ii) Advanced, on-plot piped and flush to septic tank / sewer	per life year saved (ACER): (i) 3,145 (ii) 956	none

B. List of low- and middle-income countries in 2018

Included studies must have taken place in a low-income, lower-middle income, or upper-middle income country, as classified by the World Bank in 2020. (https://datahelpdesk.worldbank.org/knowledgebase/articles/906519)

The following countries were low-income (with per capita GDP of $1,025 or less) in 2018:

Afghanistan	Guinea-Bissau	Sierra Leone
Benin	Haiti	Somalia
Burkina Faso	Korea, Dem. People's Rep.	South Sudan
Burundi	Liberia	Syrian Arab Republic
Central African Republic	Madagascar	Tajikistan
Chad	Malawi	Tanzania
Congo, Dem. Rep	Mali	Togo
Eritrea	Mozambique	Uganda
Ethiopia	Nepal	Yemen, Rep.
Gambia, The	Niger	
Guinea	Rwanda	

The following countries were lower middle-income (with per capita GDP between $1,026 and $3,995) in 2018:

Angola	India	Papua New Guinea
Bangladesh	Indonesia	Philippines
Bhutan	Kenya	São Tomé and Principe
Bolivia	Kiribati	Senegal
Cabo Verde	Kyrgyz Republic	Solomon Islands
Cambodia	Lao PDR	Sudan
Cameroon	Lesotho	Timor-Leste
Comoros	Mauritania	Tunisia
Congo, Rep.	Micronesia, Fed. Sts.	Ukraine
Côte d'Ivoire	Moldova	Uzbekistan
Djibouti	Mongolia	Vanuatu
Egypt, Arab Rep.	Morocco	Vietnam
El Salvador	Myanmar	West Bank and Gaza
Eswatini	Nicaragua	Zambia
Ghana	Nigeria	Zimbabwe
Honduras	Pakistan	

The following countries were upper middle-income (with per capita GDP between $3,996 and $12,375) in 2018:

Albania	Fiji	Namibia
Algeria	Gabon	Nauru
American Samoa	Georgia	North Macedonia
Argentina	Grenada	Paraguay
Armenia	Guatemala	Peru
Azerbaijan	Guyana	Romania
Belarus	Iran, Islamic Rep.	Russian Federation
Belize	Iraq	Samoa
Bosnia and Herzegovina	Jamaica	Serbia
Botswana	Jordan	Sri Lanka
Brazil	Kazakhstan	South Africa
Bulgaria	Kosovo	St. Lucia
China	Lebanon	St. Vincent and the Grenadines
Colombia	Libya	Suriname
Costa Rica	Malaysia	Thailand
Cuba	Maldives	Tonga
Dominica	Marshall Islands	Turkey
Dominican Republic	Mauritius	Turkmenistan
Equatorial Guinea	Mexico	Tuvalu
Ecuador	Montenegro	Venezuela, RB

C. Search strategy

Search terms for other OVID-SP and Web of Science interfaces

Table 8: Search terms for OVID-SP interface (Medline, Global Health, Embase, EconLit)

For EMBASE and Medline, we used search filters to identify economic evaluations recommended by the InterTASC Information Specialists' Sub-Group (ISSG Search Filter Resource, 2008)

	WATER QUALITY
1	(Water adj3 (treat* or quality or clean* or safe or improve* or hygiene*)).ab,ti.
2	(Water adj3 (purif* or chlor* or decontaminat* or filt* or disinfect* or floccul*)).ab,ti.
3	(Water adj3 (stor* or recontamination or re-contamination)).ab,ti.
4	Water Purification/ or Water Pollution/
5	((Drink* or consum*) adj3 water).ab,ti.
6	1 or 2 or 3 or 4
7	6 and 5
	WATER SUPPLY AND DISTRIBUTION
8	(Water adj3 (supply or availability or access or connect* or distance or improve* or distribut* or quantity or volume)).ab,ti.
9	Water Supply/
10	8 or 9
	SANITATION
11	(toilet* or latrine* or pit or pits or Sanita* or ecosan).ab,ti.
12	((feces or faeces or fecal or faecal or excre* or waste) adj3 (dispos* or manag* or service*)).ab,ti.
13	(sewage or sewer* or sewerage).ab,ti.
14	(Open adj1 defecation).ab,ti. or Sanitation/
15	11 or 12 or 13 or 14
16	7 or 10 or 15
	OUTCOMES
17	(Cost-benefit or Benefit-cost or Cost-effective* or Cost-utility or (Econ* adj1 eval*) or (cost adj2 Efficien*) or Value for Money).ab,ti.
18	16 and 17 (combines all terms)
18	limit 18 to yr="1980 -Current"

Table 9: Search terms for Web of Science

	WATER QUALITY
1	TS=(Water NEAR/3 (treat* or quality or clean* or safe or improve* or hygiene*))
2	TS=(Water NEAR/3 (purif* or chlor* or decontaminat* or filt* or disinfect* or floccul*))
3	TS= (Water NEAR/3 (stor* or recontamination or re-contamination))
4	TS=((Drink* or consum*) NEAR/3 water)
5	1 or 2 or 3
6	4 and 5
	WATER SUPPLY AND DISTRIBUTION
7	TS= (Water NEAR/3 (supply or availability or access or connect* or distance or improve* or distribut* or quantity or volume))
	SANITATION
8	TS=(toilet* or latrine* or pit or pits or Sanita* or ecosan)
9	TS=((feces or faeces or fecal or faecal or excre* or waste) NEAR/3 (dispos* or manag* or service*))
10	TS=(sewage or sewer* or sewerage)
11	TS=(Open NEAR/1 defecation)
12	8 or 9 or 10 or 11
13	6 or 7 or 12
	OUTCOMES
14	TS=(Cost-benefit or Benefit-cost or Cost-effective* or Cost-utility or (Econ* NEAR/1 eval*) or (cost adj2 Efficien*) or Value for Money)
15	13 and 14 (combines all terms)

Search terms for other databases

We also searched the below databases and websites

- International bibliography of the social sciences
- Cochrane library
- National Bureau of Economic Research
- GreenFILE (EBSCO)
- ELDIS (Institute of Development Studies)
- Centre for Environmental Health Action (World Health Organisation)
- International Initiative for Impact Evaluation
- Independent Evaluation Group (World Bank)
- IDEAS/REPec (economics working papers)
- British Library for Development Studies (Institute of Development Studies)
- JOLIS library catalogue (World Bank / International Monetary Fund)
- Research for Development (Department for International Development)

- Africa-Wide Information
- WHO regional databases (Global Index Medicus)
- Greylit.org
- Opengrey.eu
- Copenhagen Consensus Centre database

For these sources, we adapted the earlier listed search terms to their interfaces. For the simplest interfaces, our basic search strategies for water and sanitation (separately) were:

1. ("water supply" OR "drinking water") AND (cost-effective OR cost-benefit)
2. sanitation AND (cost-effective OR cost-benefit)

D. Excluded studies

As outlined in the protocol, we excluded studies if they:

- Reported a partial economic evaluation (e.g. analysis of costs or benefits alone)
- Reported a study which did not involve evaluating a drinking water or sanitation intervention as defined in the protocol
- Reported a multi-sectoral intervention for which results for the drinking water or sanitation component were not disaggregated.
- Reported a prospective economic appraisal as part of project planning
- Reported an intervention taking place in institutional settings, such as healthcare facilities, schools or work places.
- Provided insufficient detail on the intervention, methods or results, such that these could not be independently assessed.
- Were published before 1980
- Were available only in a language other than English
- Duplicated another included study
- Took place in a high-income country (HIC) rather than a LMIC as per the 2019/20 World Bank list (supplementary material B)

Studies excluded as part of full-text screening are below, tabulated by reason for exclusion.

Table 10: Studies excluded as part of full-text screening

ID	First author	Year	Reason for exclusion	Full reference
Intervention, methods or results too unclear				
1005	Yao	1982	intervention, methods or results too unclear	Yao, F. (1982). The economic aspects of the ivory coast water project: a social benefit-cost analysis. University of Southern California PhD thesis.
1028	Pushpangadan	2002	intervention, methods or results too unclear	Pushpangadan, K. (2002). Social returns from drinking water, sanitation and hygiene education : a case study of two coastal villages in Kerala. Working Paper / CDS (Kerala), no. 333(May), 29 p. : 2 fig., 8 tab.
1048	Renwick	2007	intervention, methods or results too unclear	Renwick, M., Joshi, D., Huang, M., Kong, S., Petrova, S., Bennett, G., … Jayasingh, G. (2007). Multiple Use Water Services for the Poor: Assessing the State of Knowledge. Winrock International

ID	First author	Year	Reason for exclusion	Full reference
1052	Prihandrijanti	2008	intervention, methods or results too unclear	Prihandrijanti, M., Malisie, A., & Otterpohl, R. (2008). Cost–Benefit Analysis for Centralized and Decentralized Wastewater Treatment System (Case Study in Surabaya-Indonesia).
1057	Schuen	2009	intervention, methods or results too unclear	Schuen, R., Parkinson, J., & Knapp, A. (2009). Study for financial and economic analysis of ecological sanitation in Sub-Saharan Africa, xi, 32 p. : 14 fig., 9 tab.
1063	Tang	2009	intervention, methods or results too unclear	Tang, C. (2009). Water Quality Study and Cost-Benefit Analysis of Rainwater Harvesting in. Analysis.
1078	Cameron	2011	intervention, methods or results too unclear	Cameron, J., Jagals, P., Hunter, P. R., Pedley, S., & Pond, K. (2011). Economic assessments of small-scale drinking-water interventions in pursuit of MDG target 7C. Science of the Total Environment, 410–411, 8–15. https://doi.org/10.1016/j.scitotenv.2011.09.054
1080	Uneze	2012	intervention, methods or results too unclear	Uneze, E., Tajudeen, I., & Iweala, O. (2012). Cost-effectiveness and benefit-cost analyses of some water interventions in Nigeria: the case of Bauchi State. JOURNAL OF DEVELOPMENT EFFECTIVENESS, 4(4), 497–514. https://doi.org/10.1080/19439342.2012.716075
1083	Sher	2012	intervention, methods or results too unclear	Sher, M. A. A., Hasnain, S. A., & Iqbal, S. R. (2012). Socio-economic economic analysis of the interventions aimed at improving water and sanitation condition of rural community. Research on Humanities and Social Sciences, 2(10), 57–63.
1096	Sardar	2013	intervention, methods or results too unclear	Sardar, T., Mukhopadhyay, S., Bhowmick, A. R., Chattopadhyay, J., T., S., S., M., & A.R., B. (2013). An optimal cost effectiveness study on Zimbabwe cholera seasonal data from 2008-2011. PLoS ONE, 8(12), e81231. https://doi.org/http://dx.doi.org/10.1371/journal.pone.0081231
1100	Ren	2013	intervention, methods or results too unclear	Ren, D. J., Colosi, L. M., & Smith, J. A. (2013). Evaluating the sustainability of ceramic filters for point-of-use drinking water treatment. Environmental Science & Technology, 47(19), 11206–11213. https://doi.org/http://dx.doi.org/10.1021/es4026084
1101	Molinos-Senante	2014	intervention, methods or results too unclear	Molinos-Senante, M., Carrera, A. P., Hernandez-Sancho, F., Fernandez-Cirelli, A., & Sala-Garrido, R. (2014). Economic feasibility study for improving drinking water quality: a case study of arsenic contamination in rural Argentina. EcoHealth, 11(4), 476–490. https://doi.org/http://dx.doi.org/10.1007/s10393-014-0948-5
1102	Terryn	2014	intervention, methods or results too unclear	Terryn, I. C. C., Lazar, I., Nedeff, V., & Lazar, G. (2014). Conventional vs. vacuum sewerage system in rural areas - an economic and environmental approach. Environmental Engineering and Management Journal, 13(8), 1847–1859.

ID	First author	Year	Reason for exclusion	Full reference
1104	Sepe	2014	intervention, methods or results too unclear	Sepe, M. B., Sagadal, J. N., Lange, R. D., & Porras, J. C. (2014). Biosand water filter and poor households in the Philippines. Bio-Innovation and Poverty Alleviation: Case Studies from Asia, 3–25.
1126	Azqueta	2017	intervention, methods or results too unclear	Azqueta, D., & Montoya, A. (2017). The social benefits of water and sanitation projects in Northern Colombia: Cost-Benefit Analysis, the Water Poverty Index and beyond. DEVELOPMENT POLICY REVIEW, 35(2), O118–O139. https://doi.org/10.1111/dpr.12285
1134	Weis	2017	intervention, methods or results too unclear	Weis, D., Hutton, G., & Kumar, M. (2017). The National Rural Drinking Water Security Pilot Project in India: A Cost-Benefit Analysis. WATER ECONOMICS AND POLICY, 3(4). https://doi.org/10.1142/S2382624X17500096
1148	Limantara	2018	intervention, methods or results too unclear	Limantara, L. M., Suhardjono, Rispiningtati, Fidari, J. S., & Novitasari, S. (2018). Water economic value of fresh water system in the tanggunggunung village, indonesia. INTERNATIONAL JOURNAL OF GEOMATE, 15(50), 113–120. https://doi.org/10.21660/2018.50.46457
1163	Weis	2019	intervention, methods or results too unclear	Weis, D., Hutton, G., & Kumar, M. (2018). Health costs and benefits from a pilot rural sanitation intervention in India. Journal of Water, Sanitation and Hygiene for Development, (November), 1–10. https://doi.org/10.2166/washdev.2018.076
1164	Mohapatra	2019	intervention, methods or results too unclear	Mohapatra, D. R. (2019). An economic analysis of urban infrastructure development project in the northern Indian city of Bharatpur. GROWTH AND CHANGE, 50(3), 1134–1163. https://doi.org/10.1111/grow.12299
1166	Adesina	2019	intervention, methods or results too unclear	Adesina, P. P. (2019). Integrated Investment Appraisal of Water and Sanitation Projects : A Case of Senegal Water and Sanitation Project.
1171	Ibeje	2019	intervention, methods or results too unclear	Ibeje, A. O. (2019). Cost- Benefit Analysis of Urban Water Supply and Distribution Scheme. J. Appl. Sci. Environ. Manage., 23(February), 365–370.
Partial economic evaluation				
1016	Perkins	1994	partial economic evaluation	Perkins, F. (1994). Cost Effectiveness of Water Supply Technologies in Rural Indonesia: Evidence from Nusa Tenggara Barat. Bulletin of Indonesian Economic Studies, 30(2), 91–117. https://doi.org/10.1080/00074919412331336607
1021	El Geriani	1998	partial economic evaluation	El Geriani, A. M., Essamin, O., Gijsbers, P. J. A., & Loucks, D. P. (1998). Cost-effectiveness analyses of Libya’s water supply system. JOURNAL OF WATER RESOURCES PLANNING AND MANAGEMENT-ASCE, 124(6), 320–329. https://doi.org/10.1061/(ASCE)0733-9496(1998)124:6(320)

ID	First author	Year	Reason for exclusion	Full reference
1035	Soto Montes de Oca	2005	partial economic evaluation	Soto Montes de Oca, G., & Bateman, I. J. (2005). Cost-Benefit Analysis of Urban Water Supply in Mexico City, 361–380.
1038	Oca	2006	partial economic evaluation	Oca, G. S. M. de, & Bateman, I. J. (2006). Scope sensitivity in households' willingness to pay for maintained and improved water supplies in a developing world urban area: investigating the influence of baseline supply quality and income distribution upon stated preferences in Mexico City. Water Resources Research, 42(7), W07421. https://doi.org/http://dx.doi.org/10.1029/2005WR003981
1042	Ghisi	2007	partial economic evaluation	Ghisi, E., & Ferreira, D. F. (2007). Potential for potable water savings by using rainwater and greywater in a multi-storey residential building in southern Brazil. BUILDING AND ENVIRONMENT, 42(7), 2512–2522. https://doi.org/10.1016/j.buildenv.2006.07.019
1049	Roy	2007	partial economic evaluation	Roy, J. (2007). Estimating the Economic Benefits of Arsenic Removal in India: A Case Study from West Bengal. Working Papers, (21).
1060	Sijbesma	2009	partial economic evaluation	Sijbesma, C., & Christoffers, T. (2009). The value of hygiene promotion: Cost-effectiveness analysis of interventions in developing countries. Health Policy and Planning, 24(6), 418–427. https://doi.org/10.1093/heapol/czp036
1186	Pattanayak	2010	partial economic evaluation	Pattanayak, S. K., Poulos, C., Yang, J.-C., & Patli, S. (2010). How valuable are environmental health interventions? Evaluation of water and sanitation programmes in India. BULLETIN OF THE WORLD HEALTH ORGANIZATION, 88(7), 535–542. https://doi.org/10.2471/BLT.09.066050
1072	Kremer	2011	partial economic evaluation	Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2011). Spring cleaning: Rural water impacts, valuation, and property rights institutions. Quarterly Journal of Economics, 126(1), 145–205. https://doi.org/10.1093/qje/qjq010
1076	Horn	2011	partial economic evaluation	Horn, T. (2008). Welfare Effects of Access to Water Service in Cambodia. Discussion Papers In Economics And Business (Vol. 11–08).
1115	Rocha	2015	partial economic evaluation	Rocha, R., & Soares, R. (2015). Water scarcity and birth outcomes in the Brazilian semiarid. Journal of Development Economics, 112, 72–91. https://doi.org/http://dx.doi.org/10.1016/j.jdeveco.2014.10.003
1116	Briceno	2015	partial economic evaluation	Briceño, B., & Chase, C. (2015). Cost-efficiency of rural sanitation promotion : activity-based costing and experimental evidence from Tanzania. Journal of Development Effectiveness, 7(4), 423–434. https://doi.org/10.1080/19439342.2015.1105848

ID	First author	Year	Reason for exclusion	Full reference
1120	Vishwakarma	2016	partial economic evaluation	Vishwakarma, A., Kulshrestha, M., Amulya Nyathikala, S., & Kulshreshtha, M. (2016). Cost efficiency benchmarking of urban water supply utilities: The case of an Indian state. Water and Environment Journal, 30(1–2), 77–87. https://doi.org/10.1111/wej.12171
1129	Nandi	2017	partial economic evaluation	Nandi, A., Megiddo, I., Ashok, A., Verma, A., & Laxminarayan, R. (2017). Reduced burden of childhood diarrheal diseases through increased access to water and sanitation in India: A modeling analysis. Social Science and Medicine, 180, 181–192. https://doi.org/10.1016/j.socscimed.2016.08.049
1139	Reygadas	2018	partial economic evaluation	Reygadas, F., Gruber, J. S., Dreizler, L., Nelson, K. L., & Ray, I. (2018). Measuring user compliance and cost effectiveness of safe drinking water programs: A cluster-randomized study of household ultraviolet disinfection in Rural Mexico. American Journal of Tropical Medicine and Hygiene, 98(3), 824–834. https://doi.org/10.4269/ajtmh.17-0440
1146	Bassi	2018	partial economic evaluation	Bassi, S. A., Tange, I., Holm, B., Boldrin, A., & Rygaard, M. (2018). A multi-criteria assessment of water supply in Ugandan refugee settlements. Water, 10(10), 1493. https://doi.org/http://dx.doi.org/10.3390/w10101493
1152	Creaco,	2018	partial economic evaluation	Creaco, E., & Walski, T. (2018). Operation and Cost-Effectiveness of Local and Remote RTC. JOURNAL OF WATER RESOURCES PLANNING AND MANAGEMENT, 144(11). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000993
1153	Tudela-Mamani	2018	partial economic evaluation	Tudela-Mamani, J. W., Leos-Rodriguez, J. A., & Zavala-Pineda, M. J. (2018). Estimation of economic benefits for improvements in basic sanitation services using the Contingent Valuation Method. Agrociencia (Montecillo), 52(3), 467–481.
1154	Woode	2018	partial economic evaluation	Woode, P. K., Dwumfour-Asare, B., Nyarko, K. B., & Appiah-Effah, E. (2018). Cost and effectiveness of water, sanitation and hygiene promotion intervention in Ghana: the case of four communities in the Brong Ahafo region. Heliyon, 4(10), e00841. https://doi.org/http://dx.doi.org/10.1016/j.heliyon.2018.e00841
1155	Laramee	2018	partial economic evaluation	Laramee, J., Tilmans, S., & Davis, J. (2018). Costs and benefits of biogas recovery from communal anaerobic digesters treating domestic wastewater: Evidence from peri-urban Zambia. Journal of Environmental Management, 210, 23–35. https://doi.org/10.1016/j.jenvman.2017.12.064
1168	Augsburg	2019	partial economic evaluation	Augsburg, B., Caeyers, B., & Malde, B. (2019). Can micro-credit support public health subsidy programs? Policy Research Working Paper - World Bank, (8846), 48-pp.

ID	First author	Year	Reason for exclusion	Full reference
Not an economic evaluation				
1004	Cvjetanović	1981	not an economic evaluation	Cvjetanović, B. (1981). The Cost-Effectiveness of Alternative Measures for Control of Enteric Diseases. In *The Impact of Interventions in Water Supply and Sanitation in Developing Countries: Proceedings of a Seminar held at the Pan American Health Organization March 25-26, 1980, Washington, D.C.* (pp. 81–90). https://doi.org/10.1088/0953-2048/24/4/045011
1006	Feachem	1984	not an economic evaluation	Feachem, R. G. (1984). Interventions for the control of diarrhoeal diseases among young children: Promotion of personal and domestic hygiene. Bulletin of the World Health Organization, 62(3), 467–476.
1007	Briscoe	1984	not an economic evaluation	Briscoe, J. (1984). Water supply and health in developing countries: selective primary health care revisited. American Journal of Public Health, 74(9), 1009–1013.
1009	Briscoe	1985	not an economic evaluation	Briscoe, J. (1985). Evaluating water supply and other health programs: Short-run vs long-run mortality effects. Public Health, 99(3), 142–145. https://doi.org/10.1016/S0033-3506(85)80103-7
1010	Esrey	1986	not an economic evaluation	Esrey, S. A., & Habicht, J. P. (1986). Epidemiologic evidence for health benefits from improved water and sanitation in developing countries. Epidemiologic Reviews, 8, 117–128.
1011	Feachem	1986	not an economic evaluation	Feachem, R. G. (1986). Preventing diarrhoea: what are the policy options? Health Policy and Planning, 1(2), 109–117. https://doi.org/10.1093/heapol/1.2.109
1012	Briscoe	1987	not an economic evaluation	Briscoe, J. (1987). A role for water supply and sanitation in the child survival revolution. Bulletin of the Pan American Health Organization, 21(2), 93–105.
1013	MacRae	1988	not an economic evaluation	MacRae Jr, D., & Whittington, D. (1988). Assessing Preferences in Cost-Benefit Analysis: Reflections on Rural Water Supply Evaluation in Haiti. Journal of Policy Analysis and Management, 7(2), 246–263.
1015	WRC	1991	not an economic evaluation	WRC. (1991). Guidelines on the cost effectiveness of rural water supply and sanitation projects.
1017	Kirkpatrick	1996	not an economic evaluation	Kirkpatrick, C. H., & Weiss, J. (1996). Cost-benefit Analysis and Project Appraisal in Developing Countries, 25, 321.
1018	Chatterji	1997	not an economic evaluation	Chatterji, D. (1997). Economic evaluation of urban water supply schemes. ECONOMIC AND POLITICAL WEEKLY, 32(26), 1542–1546.

ID	First author	Year	Reason for exclusion	Full reference
1019	Joubert	1997	not an economic evaluation	Joubert, A. R., Leiman, A., DeKlerk, H. M., Katua, S., Aggenbach, L. C., & Joubert et al, A. R. (1997). Fynbos (fine bush) vegetation and the supply of water: a comparison of multi-criteria decision analysis and cost-benefit analysis. ECOLOGICAL ECONOMICS, 22(2), 123–140. https://doi.org/10.1016/S0921-8009(97)00573-9
1023	Tao	1999	not an economic evaluation	Tao, Y. X., & Hills, P. (1999). Assessment of alternative wastewater treatment approaches in Guangzhou, China. Water Science and Technology, 39(5), 227–234. https://doi.org/10.1016/S0273-1223(99)00106-7
1024	Clarke	2000	not an economic evaluation	Clarke, G., Menard, C., & Zuluaga, A. M. (1999). The Welfare Effects of Private Sector Participation in Guinea’s Urban Water Supply.
1025	Russell	2001	not an economic evaluation	Russell, C. S. (2001). Investing in water quality: Measuring benefits, costs and risks.
1033	Rijsberman	2004	not an economic evaluation	Rijsberman, F. (2004). Sanitation and access to clean water. In B. Lomborg (Ed.), Global Crises, Global Solutions (pp. 498–527). https://doi.org/10.1017/CBO9780511492624.010
1040	Whittington	2006	not an economic evaluation	Whittington, D., & Hanemann, W. M. (2006). The Economic Costs and Benefits of Investments in Municipal Water and Sanitation Infrastructure : A Global Perspective (No. 1027). CUDARE Working Papers.
1050	Hutton	2008	not an economic evaluation	Hutton, G. (2008). Economic evaluation of environmental health interventions to support decision making. Environmental Health Insights, 2, 137–155.
1062	Dhaliwal	2009	not an economic evaluation	Dhaliwal, I., & Tulloch, C. (2009). Cost-effectiveness of interventions to prevent child diarrhea (Abdul Latif Jameel Poverty Action Lab). Cambridge, MA.
1064	Agoramoorthy	2009	not an economic evaluation	Agoramoorthy, G., & Hsu, M. J. (2009). India needs sanitation policy reform to enhance public health. Journal of Economic Policy Reform, 12(4), 333–342. https://doi.org/10.1080/17487870903314625
1065	Carrard	2010	not an economic evaluation	Carrard, N., Willetts, J., Mitchell, C., Paddon, M., & Retamal, M. (2010). Selecting Sanitation Solutions for Peri-urban Areas: A Case Study of Can Tho, Vietnam. Water Practice and Technology, 5(4), 1–13. https://doi.org/10.2166/wpt.2010.109
1067	Ahuja	2010	not an economic evaluation	Ahuja, A., Kremer, M., & Zwane, A. P. (2010). Providing safe water: evidence from randomized evaluations. Annual Review of Resource Economics, 2, 237–256.

ID	First author	Year	Reason for exclusion	Full reference
1068	Willetts	2010	not an economic evaluation	Willetts, J., Carrard, N., Mitchell, C., Trung, N. H., Nam, N. D. G., & Paddon, M. (2010). Cost-effectiveness analysis as a methodology to compare sanitation options in peri-urban Can Tho , Vietnam. In Pumps, Pipes and Promises - A collection of papers from the IRC Symposium 2010 (pp. 144–159).
1074	Rashid	2011	not an economic evaluation	Rashid, M. M., & Hayes, D. F. (2011). Needs-based sewerage prioritization: Alternative to conventional cost-benefit analysis. Journal of Environmental Management, 92(10), 2427–2440. https://doi.org/10.1016/j.jenvman.2011.05.002
1075	Kumar	2011	not an economic evaluation	Kumar, S. G., Kar, S. S., & Jain, A. (2011). Health and environmental sanitation in India: issues for prioritizing control strategies. Indian Journal of Occupational and Environmental Medicine, 15(3), 93–96. https://doi.org/http://dx.doi.org/10.4103/0019-5278.93196
1077	Cameron	2011	not an economic evaluation	Cameron, J., Hunter, P., Jagals, P., & Pond, K. (2011). Valuing water, valuing livelihoods: guidance on social cost-benefit analysis of drinking-water interventions, with special reference to small community water supplies. London.
1079	Hunt	2011	not an economic evaluation	Hunt, A. (2011). Policy Interventions to Address Health Impacts Associated with Air Pollution , Unsafe Water Supply and Sanitation , and Hazardous Chemicals.
1086	Harikumar	2012	not an economic evaluation	Harikumar, P. S., & Bindhya, M. K. (2012). A synoptic study on the preparation of a liquid waste management plan for Kerala State, India. Environment and Natural Resources Research, 2(2), 74–83.
1188	Rijsberman	2012	not an economic evaluation	Rijsberman, F., & Zwane, A. P. (2012). Copenhagen Consensus 2012: Sanitation and Water Challenge Paper.
1098	Basu	2013	not an economic evaluation	Basu, N. B., Dey, A., & Ghosh, D. (2013). Kolkata's brick sewer renewal: history, challenges and benefits. PROCEEDINGS OF THE INSTITUTION OF CIVIL ENGINEERS-CIVIL ENGINEERING, 166(2), 74–81. https://doi.org/10.1680/cien.12.00016
1103	Novak	2014	not an economic evaluation	Novak, L. (2014). The impact of access to water on child health in Senegal. Review of Development Economics, 18(3), 431–444. https://doi.org/http://dx.doi.org/10.1111/rode.12094
1110	Yates	2015	not an economic evaluation	Yates, T., Lantagne, D., Mintz, E., & Quick, R. (2015). The impact of water, sanitation, and hygiene interventions on the health and well-being of people living with HIV: A systematic review. Journal of Acquired Immune Deficiency Syndromes, 68(Suppl. 3), S318–S330. https://doi.org/10.1097/QAI.0000000000000487

ID	First author	Year	Reason for exclusion	Full reference
1111	Behnsen	2015	not an economic evaluation	Behnsen, F., Cenko, P., Shehu, K., & Ndreu, P. (2015). Energy efficient water supply systems in Albania: moving from energy scans and cost-benefit analysis to the planning, implementation and management of energy efficiency water infrastructure in Albania. Water Asset Management International, 11(3), 8–12.
1114	Hopkins	2015	not an economic evaluation	Hopkins, O. S. (2015). A regional approach to optimizing the location of rural handpumps. JOURNAL OF WATER SANITATION AND HYGIENE FOR DEVELOPMENT, 5(3), 493–501. https://doi.org/10.2166/washdev.2015.128
1117	Shekar	2016	not an economic evaluation	Shekar, M., Dayton Eberwein, J., & Kakietek, J. (2016). The costs of stunting in South Asia and the benefits of public investments in nutrition. Maternal and Child Nutrition, 12, 186–195. https://doi.org/10.1111/mcn.12281
1124	Zhang	2016	not an economic evaluation	Zhang, J., & Xu, L. C. (2016). The Long-Run Effects of Treated Water on Education: The Rural Drinking Water Program in China. Journal of Development Economics, 122(0), 1–15. https://doi.org/10.1016/j.jdeveco.2016.04.004
1130	Mock	2017	not an economic evaluation	Mock, C. N., Smith, K. R., Kobusingye, O., Nugent, R., Abdalla, S., Ahuja, R. B., … Watkins, D. A. (2017). Injury Prevention and Environmental Health : Key Messages from Disease Control Priorities , Third Edition, 1–23.
1144	Bernal	2018	not an economic evaluation	Bernal, D. (2018). A conceptual model for decentralized municipal wastewater management. Water Practice & Technology, 13(1), 134–142. https://doi.org/http://dx.doi.org/10.2166/wpt.2018.022
1161	Berhe	2019	not an economic evaluation	Berhe, H. W., Makinde, O. D., & Theuri, D. M. (2019). Co-dynamics of measles and dysentery diarrhea diseases with optimal control and cost-effectiveness analysis. APPLIED MATHEMATICS AND COMPUTATION, 347, 903–921. https://doi.org/10.1016/j.amc.2018.11.049
1165	Cui	2019	not an economic evaluation	Cui, C., Wang, J., Liu, Y., & Coffey, V. (2019). Relationships among Value-for-Money Drivers of Public-Private Partnership Infrastructure Projects. JOURNAL OF INFRASTRUCTURE SYSTEMS, 25(2). https://doi.org/10.1061/(ASCE)IS.1943-555X.0000479
1169	Nauges	2019	not an economic evaluation	Nauges, C., & Whittington, D. (2019). Social norms information treatments in the municipal water supply sector: Some new insights on benefits and costs. Water Economics and Policy, 5(3). https://doi.org/10.1142/S2382624X18500261
1170	Frempong	2019	not an economic evaluation	Frey, R. L., & Frempong, R. (2019). A micro-based approach to evaluate the effect of water supply on health in Uganda, (2012).

ID	First author	Year	Reason for exclusion	Full reference
1187	Oyebode	2019	not an economic evaluation	Oyebode, O. J., & Ige, M. M. (2019). Strategic Evaluation of Cost and Benefit Analysis of a Rural Water Supply Project: A Case Study of Dei-Dei Community in Abuja, Nigeria. Journal of Water Resource Engineering and Management, (July), 11–28.
1177	Mukhopadhyay	2020	not an economic evaluation	Mukhopadhyay, B. (2020). Network Externalities and Sanitation. The Journal of Developing Areas, 54(3). https://doi.org/10.1353/jda.2020.0027
Not drinking water or sanitation				
1031	Whittington	2004	not drinking water or sanitation (water resources management)	Whittington, D., Lauria, D. T., Prabhu, V., & Cook, J. (2004). An Economic Reappraisal of the Melamchi Water Supply Project--Kathmandu, Nepal. Portuguese Economic Journal, 3(2), 157–178.
1034	Yamout	2005	not drinking water or sanitation (water resources management)	Yamout, G., & El-Fadel, M. (2005). An optimization approach for multi-sectoral water supply management in the Greater Beirut Area. WATER RESOURCES MANAGEMENT, 19(6), 791–812. https://doi.org/10.1007/s11269-005-3280-6
1051	Liang	2008	not drinking water or sanitation (wastewater reuse or rainwater harvesting for non-drinking purposes)	Liang, X., & Dijk, M. P. Van. (2008). Economic and Financial Analysis of Decentralized Water Recycling Systems in Beijing. In 3rd SWITHC Scientific Meeting (Vol. 2010, p. 15).
1058	Kannan	2009	not drinking water or sanitation (wastewater treatment technology, not user-facing)	Kannan, S., Singal, S., Kazmi, A., & Sharma, M. (2009). Selection of Appropriate Sewage Treatment Technology for Kancheepuram City. Asian Journal of Water, Environment and Pollution, 6(2), 107–111.
1061	Chen	2009	not drinking water or sanitation (wastewater treatment technology, not user-facing)	Chen, R., & Wang, X. C. (2009). Cost-benefit evaluation of a decentralized water system for wastewater reuse and environmental protection. Water Science and Technology, 59(8), 1515–1522. https://doi.org/10.2166/wst.2009.156

ID	First author	Year	Reason for exclusion	Full reference
1066	Seidu	2010	not drinking water or sanitation (wastewater treatment technology, not user-facing)	Seidu, R. (2010). Disentangling the risk factors and health risks associated with faecal sludge and wastewater reuse in Ghana. Philosophiae Doctor (PhD) Thesis 2010:17, 72–75.
1121	Agusdinata	2016	not drinking water or sanitation (agriculture)	Agusdinata, D. B. (2016). Evaluating Water Infrastructure and Agriculture Practices for Drought Adaptations in East Africa A Combined Hydrological and System Dynamics Approach. In PROCEEDINGS OF THE SIXTH IEEE GLOBAL HUMANITARIAN TECHNOLOGY CONFERENCE GHTC 2016 (pp. 753–760).
1125	Thomas	2016	not drinking water or sanitation (agriculture)	Thomas, E. A. (2016). Broken Pumps and Promises: Incentivizing Impact in Environmental Health. New York:
1118	Bai	2016	not drinking water or sanitation (water resources management)	Bai, H., Lian, J., & Zhu, Q. (2016). Emergency Water Conveyance Project into Dalian Cost Benefit Analysis. In 2016 6TH INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY FOR MANUFACTURING SYSTEMS (ITMS 2016) (pp. 313–316).
1119	Kerstens	2016	not drinking water or sanitation (water resources management)	Kerstens, S. M., Hutton, G., Firmansyah, I., Leusbrock, I., & Zeeman, G. (2016). An integrated approach to evaluate benefits and costs of wastewater and solid waste management to improve the living environment: the Citarum River in West Java, Indonesia. Journal of Environmental Protection, 7(11), 1439–1465. https://doi.org/http://dx.doi.org/10.4236/jep.2016.711122
1138	Dadhich	2017	not drinking water or sanitation (agriculture)	Dadhich, G., & Shrivastava, V. (2017). Economic Comparison of Solar PV and Diesel Water Pumping System. In 2017 IEEE INTERNATIONAL CONFERENCE ON INFORMATION, COMMUNICATION, INSTRUMENTATION AND CONTROL (ICICIC).
1127	Dominguez	2017	not drinking water or sanitation (wastewater reuse or rainwater harvesting for non-drinking purposes)	Dominguez, I., Ward, S., Mendoza, J. G., Rincon, C. I., & Oviedo-Ocana, E. R. (2017). End-user cost-benefit prioritization for selecting rainwater harvesting and greywater reuse in social housing. Water, 9(7), 516. https://doi.org/http://dx.doi.org/10.3390/w9070516

ID	First author	Year	Reason for exclusion	Full reference
1135	Cardona	2017	not drinking water or sanitation (wastewater reuse or rainwater harvesting for non-drinking purposes)	Cardona, J. A., Segovia, O. C., Bottger, S., Castillo, N. A. M., Cavallo, L., Ribeiro, I. E., & Schluter, S. (2017). Reuse-oriented decentralized wastewater and sewage sludge treatment for rural settlements in Brazil: a cost-benefit analysis. Desalination and Water Treatment, 91, 82–92. https://doi.org/http://dx.doi.org/10.5004/dwt.2017.21421
1142	Amos	2018	not drinking water or sanitation (wastewater reuse or rainwater harvesting for non-drinking purposes)	Amos, C. C., Rahman, A., & Gathenya, J. M. (2018). Economic analysis of rainwater harvesting systems comparing developing and developed countries: A case study of Australia and Kenya. JOURNAL OF CLEANER PRODUCTION, 172, 196–207. https://doi.org/10.1016/j.jclepro.2017.10.114
1141	Zhang	2018	not drinking water or sanitation (wastewater treatment technology, not user-facing)	Zhang, J., Qiu, Z., Li, F., Du, Q., JingDong, Z., ZhenZhen, Q., … Qian, D. (2018). An exploration of comprehensive evaluation method of sewage treatment construction project in small and medium towns: theory and application. DESALINATION AND WATER TREATMENT, 118, 70–78. https://doi.org/http://dx.doi.org/10.5004/dwt.2018.22418
1147	De Risi	2018	not drinking water or sanitation (water resources management)	De Risi, R., De Paola, F., Turpie, J., & Kroeger, T. (2018). Life Cycle Cost and Return on Investment as complementary decision variables for urban flood risk management in developing countries. INTERNATIONAL JOURNAL OF DISASTER RISK REDUCTION, 28, 88–106. https://doi.org/10.1016/j.ijdrr.2018.02.026
1160	Banerjee	2019	not drinking water or sanitation (agriculture)	Banerjee, O., Cicowiez, M., Horridge, M., & Vargas, R. (2019). Evaluating synergies and trade-offs in achieving the SDGs of zero hunger and clean water and sanitation: An application of the IEEM Platform to Guatemala. ECOLOGICAL ECONOMICS, 161, 280–291. https://doi.org/10.1016/j.ecolecon.2019.04.003
1156	Hallegatte	2019	not drinking water or sanitation (wastewater reuse or rainwater harvesting for non-drinking purposes)	Hallegatte, S., Rozenberg, J., Maruyama Rentschler, J. E., Nicolas, C. M., & Fox, C. J. E. (2019). Strengthening New Infrastructure Assets : A Cost-Benefit Analysis.

ID	First author	Year	Reason for exclusion	Full reference
1157	Potgieter	2019	not drinking water or sanitation (water resources management)	Potgieter, J. C., Herold, C., van Dijk, M., & Bhagwan, J. N. (2019). Economic benefit of ensuring uninterrupted water supply during prolonged electricity disruptions - City of Tshwane case study. JOURNAL OF THE SOUTH AFRICAN INSTITUTION OF CIVIL ENGINEERING, 61(4), 19–28. https://doi.org/10.17159/2309-8775/2019/v61n4a2
1172	Hooper	2020	not drinking water or sanitation (hygiene)	Hooper, A., Loy, E., Clayton, F., Nabata, K., & Riedlinger, R. (2020). Long-term feasibility, sustainability, and cost-benefit analysis of reusable menstrual hygiene management products in a population of adolescent females in a remote Himalayan community. Journal of Investigative Medicine, 68(1), A176. https://doi.org/http://dx.doi.org/10.1136/jim-2019-WMRC.406
Not in English				
1195	Lou	1990	not in English	Lou, H. (1990). Effectiveness evaluation and cost-effectiveness estimate for diarrhea control by environment improvement in rural area. Zhonghua Liu Xing Bing Xue Za Zhi, 11(3), 170–174.
1193	Murgueytio	1995	not in English	Murgueytio, P., & Estupinan-Day, S. (1995). Evaluación de costos y beneficios anticipados del Programa de Fluoruración del Agua Potable propuesto para la VIII Region, Chile.
1194	Martins	1995	not in English	Martins, G. (1995). Benefícios e custos do abastecimento de água e esgotamento sanitário em pequenas comunidades (MSc thesis). Universidade de Säo Paulo.
1192	Silva	1996	not in English	Silva, W. L. M. (1996). Subsídios para o estudo da relação custos-efetividade de siltemas de tratamento de esgoto. Escola Nacional de Saúde Pública.
1196	Xiao	1997	not in English	Xiao, S., Lin, C., & Chen, K. (1997). Evaluation of effectiveness of comprehensive control for diarrhea diseases in rural areas of east Fujian and analysis of its cost-benefit. Zhonghua yu fang yi xue za zhi [Chinese journal of preventive medicine], 31(1), 40–41.
1090	Arbage Lobo	2013	not in English	Arbage Lobo, M. A., de Lima, D. M., Nobre Souza, C. M., Nascimento, W. A., Cardoso Arajo, L. C., & dos Santos, N. B. (2013). Economic evaluation of social technologies applied to health promotion: water supply by the SODIS System in riverside communities of the Brazilian Amazon. CIENCIA & SAUDE COLETIVA, 18(7), 2119–2127. https://doi.org/10.1590/S1413-81232013000700027
1091	Gil	2013	not in English	Gil, H. A., Cisneros, J. M., Prada, J. D. de, Plevich, J. O., & Delgado, A. R. S. (2013). Green technologies for the use of urban wastewater: economic analysis. Tecnologias Verdes Para El Aprovechamiento de Aguas Residuales Urbanas: Analisis Economico., 8(3), 118–128.
Not in LMIC				

ID	First author	Year	Reason for exclusion	Full reference
1026	Licsko	2002	not in LMIC	Licskó, I., Melicz, Z., & Szabó, A. (2002). Chemical Pretreatment of Sewage -- a Cost-Benefit Method for Upgrading Existing and Constructing New Wastewater Treatment Plants. In J. Rubin, H and Nachtnebel, P and Shamir, U and Furst (Ed.), Water Resources Quality (pp. 371–389). https://doi.org/10.1007/978-3-642-56013-2_22
1027	Hauger	2002	not in LMIC	Hauger, M. B., Rauche, W., Linde, J. J., & Mikkelsen, P. S. (2002). Cost benefit risk - A concept for management of integrated urban wastewater systems? Water Science and Technology, 45(3), 185–193. https://doi.org/10.2166/wst.2002.0078
1056	Hunter	2009	not in LMIC	Hunter, P. R., Pond, K., Jagals, P., & Cameron, J. (2009). An assessment of the costs and benefits of interventions aimed at improving rural community water supplies in developed countries. Science of the Total Environment, 407(12), 3681–3685. https://doi.org/10.1016/j.scitotenv.2009.03.013
1073	Lindhe	2011	not in LMIC	Lindhe, A., Rosen, L., Norberg, T., Bergstedt, O., & Pettersson, T. J. R. (2011). Cost-effectiveness analysis of risk-reduction measures to reach water safety targets. Water Research (Oxford), 45(1), 241–253. https://doi.org/http://dx.doi.org/10.1016/j.watres.2010.07.048
1093	Petohleb Cerneha	2013	not in LMIC	Petohleb Cerneha, S., Klun, M., & Devjak, S. (2013). The Social Cost-Benefit Analysis as Estimation Methodology: Case Study for Infrastructure Projects. Mednarodna Revija Za Javno Upravo/International Public Administration Review, 11(2), 57–74.
1095	Listowski	2013	not in LMIC	Listowski, A., Ngo, H. H., & Guo, W. S. (2013). Establishment of an economic evaluation model for urban recycled water. Resources, Conservation and Recycling, 72, 67–75. https://doi.org/10.1016/j.resconrec.2012.12.011
1107	Cutler	2014	not in LMIC	Cutler, D., & Miller, G. (2014). The Role of Public Health Improvements in Health Advances: The 20th Century United States.
1123	Karolinczak	2016	not in LMIC	Karolinczak, B., & Milaszewski, R. (2016). Application of assessment methods of the economic effectiveness of water supply and sewerage facilities. ROCZNIK OCHRONA SRODOWISKA, 18(2), 770–782.
1133	Cory	2017	not in LMIC	Cory, D. C., & Taylor, L. D. (2017). On the Distributional Implications of Safe Drinking Water Standards. JOURNAL OF BENEFIT-COST ANALYSIS, 8(1), 49–90. https://doi.org/10.1017/bca.2017.2
1137	Schneider	2017	not in LMIC	Schneider, O. D., & Lechevallier, M. W. (2017). A Cost-Effective Treatment Process for Producing High-Quality Drinking Water. JOURNAL AMERICAN WATER WORKS ASSOCIATION, 109(3), 39–47. https://doi.org/10.5942/jawwa.2017.109.0026

ID	First author	Year	Reason for exclusion	Full reference
1159	Ryu	2019	not in LMIC	Ryu, J., Kim, K., Oh, M., & Shin, J. (2019). Why environmental and social benefits should be included in cost-benefit analysis of infrastructure? ENVIRONMENTAL SCIENCE AND POLLUTION RESEARCH, 26(21), 21693–21703. https://doi.org/10.1007/s11356-019-05475-6
1174	Choi	2020	not in LMIC	Choi, G. W., Jo, H. G., Park, H. S., & Jang, D. W. (2020). Application of decision making model for leakage reduction to economic project in water distribution systems. JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING. https://doi.org/10.1007/s12652-019-01634-2
Not possible to separate WASH component				
1003	Popkin	1980	not possible to separate WASH component	Popkin, B. M., Solon, F. S., Fernandez, T., & Lantham, M. C. (1980). Benefit-cost analysis in the nutrition area: A project in the Philippines. Social Science and Medicine. Part C Medical Economics, 14(3), 207–216. https://doi.org/10.1016/S0160-7995(80)80004-9
1008	Wang'ombe	1984	not possible to separate WASH component	Wang'ombe, J. K. (1984). Economic evaluation in primary health care: The case of Western Kenya community based health care project. Social Science and Medicine, 18(5), 375–385. https://doi.org/10.1016/0277-9536(84)90055-8
1094	Kern	2013	not possible to separate WASH component	Kern, E., Verguet, S., Yuhas, K., Odhiambo, F. H., Kahn, J. G., & Walson, J. (2013). Provision of bednets and water filters to delay HIV-1 progression: Cost-effectiveness analysis of a Kenyan multisite study. Tropical Medicine and International Health, 18(8), 916–924. https://doi.org/10.1111/tmi.12127
1113	Verguet	2015	not possible to separate WASH component	Behnsen, F., Cenko, P., Shehu, K., & Ndreu, P. (2015). Energy efficient water supply systems in Albania: moving from energy scans and cost-benefit analysis to the planning, implementation and management of energy efficiency water infrastructure in Albania. Water Asset Management International, 11(3), 8–12.
PDF not retrievable				
1190	Hulton	1980	PDF not retrievable	Hulton, M. R. (1980). Report on cost effectiveness appraisals of water supplies in Africa (MSc thesis). University of Newcastle upon Tyne.
1182	Dooyeweerd	1983	PDF not retrievable	Dooyeweerd, E. (1983). The use of social cost benefit analysis for rural water supply and sanitation; case study, Indramayu, West Java, Indonesia. Erasmus University.
1191	Fredriksson	1989	PDF not retrievable	Fredriksson;, P., & Persson, A. (1989). The Manicaland Health, Water and Sanitation Pragramme in Zimbabwe : a social cost benefit analysis. Stockholm School of Economics.
Another reporting of an included study				

ID	First author	Year	Reason for exclusion	Full reference
1047	Hutton	2007	another reporting of an included study	Hutton, G. (2007). Unsafe water and lack of sanitation. Solutions for the World's Biggest Problems: Costs and Benefits, 405–424. https://doi.org/10.1017/CBO9780511493560.023
1053	Whittington	2008	another reporting of an included study	Dale Whittington, Hanemann, W. M., Sadoff, C., & Jeuland, M. (2008). Copenhagen Consensus 2008 Challenge Paper: Sanitation and Water. Copenhagen Consensus Center.
1088	Kahn	2012	another reporting of an included study	Kahn, J. G., Muraguri, N., Harris, B., Lugada, E., Clasen, T., Grabowsky, M., … Shariff, S. (2012). Integrated HIV testing, malaria, and diarrhea prevention campaign in Kenya: Modeled health impact and cost-effectiveness. PLoS ONE, 7(2), 1–9. https://doi.org/10.1371/journal.pone.0031316
1106	Hutton	2014	another reporting of an included study	Hutton, G., Rodriguez, U. P., Winara, A., Viet-Anh, N., Phyrum, K., Chuan, L., … Weitz, A. (2014). Economic efficiency of sanitation interventions in Southeast Asia. Journal of Water Sanitation and Hygiene for Development, 4(1), 23–36.
1150	Hutton	2018	another reporting of an included study	Hutton, G. (2018). Global Benefits and Costs of Achieving Universal Coverage of Basic Water and Sanitation Services as part of the 2030 Agenda for Sustainable Development. In B. Lomborg (Ed.), Prioritizing Development: A Cost Benefit Analysis of the United Nations' Sustainable Development Goals (pp. 422–445). Cambridge: Cambridge University Press. https://doi.org/DOI: 10.1017/9781108233767.025
Investment case for planned programme				
1041	Renwick	2007	investment case for planned programme	Renwick, M., Subedi, P. S., & Hutton, G. (2007). Biogas for Better Life : an African Initiative a Cost -Benefit Analysis of National and Regional Integrated Biogas and Sanitation Programs in Sub -Saharan Africa.

E. Scoring protocol for CHEERS

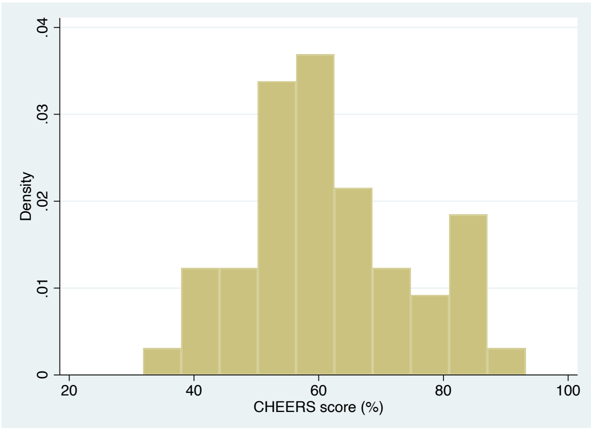
This table below sets out the protocol we developed for scoring each CHEERS reporting item.

Area	Reporting item (from Husereau et al., 2013)		Interpretation for scoring in this study (authors)		
Framing	**1. Title**	Identifies the study as an EE or uses specific terms such as CEA, and describes the interventions compared.	1 - includes, CEA, CBA, EE or similar in title	0.5 - implication of EE but not with clear terms	0 - nothing implies EE
	2. Abstract	Structured summary of objectives, perspective, setting, methods, results (including base case & uncertainty)	1 - abstract contains most of these	0.5 - abstract provides some of above, but missing some of the most important (intervention, methods, results)	0 - abstract provides very few of these
	3. Background and objectives	Provides broader context for the study. Presents the study question and its relevance for policy or practice decisions.	1 - study aim/question presented with relevance for decisions	0.5 - EITHER clear aim but limited context OR good relevance re decisions, but unclear aim or question regarding EE	0 - BOTH aim/question AND relevance for decisions unclear
Population, setting	**4. Target population and subgroups**	Describes characteristics of the base case population and subgroups analysed, including why they were chosen.	1 - study population sufficiently described, including any assumptions	0.5 - study population described, but too briefly in respect of policy-relevant characteristics	0 - study population unclear
	5. Setting and location	States relevant aspects of the system(s) in which the decision(s) need(s) to be made.	1 - characteristics of decision context described	0.5 - decision context described but insufficiently	0 - decision context appears not to have been considered
Key methods decisions	**6. Study perspective**	Describes the perspective of the study and relate this to the costs being evaluated.	1 - perspective stated clearly and interpreted	0.5 - perspective vaguely stated, but sufficient detail such that it can easily be discerned	0 - perspective unclear
	7. Comparators	Describes the interventions or strategies being compared and state why they were chosen.	1 - intervention and comparator both well-described	0.5 - limited detail on comparator, but doesn't substantially harm interpretation	0 - lack of detail on comparator limits interpretation
	8. Time horizon	States the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.	1 - time horizon stated and explained	0.5 - time horizon stated without context, or clearly implied	0 - time horizon unclear, which limits interpretation
	9. Discount rate	Reports the choice of discount rate(s) used for costs and outcomes and say why appropriate.	1 - rate reported and justified	0.5 - rate reported but unclear how applied or not justified / referenced	0 - no discounting, or unclear how applied

Area	Reporting item (from Husereau et al., 2013)		Interpretation for scoring in this study (authors)		
Outcomes and costs	**10. Choice of outcomes**	Describes what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.	1 - outcomes / benefits clear	0.5 - lack of clarity on some but not all benefits	0 - outcomes / benefits very unclear
	11. Measurement of effectiveness	*(i) Single study:* Describes features of design and why sufficient, (ii) *Synthesis-based*: describes identification & synthesis of included studies	1 - source of effects data explained, discussed and justified	0.5 - source and important assumptions clear, but not discussed / justified	0 - source or important assumptions unclear
	12. Measurement & valuation of pref.-based outcomes	If applicable, describe the population and methods used to elicit preferences for outcomes.	1 - if DALYs used, source of weights referenced / discussed	0.5 - slightly unclear on weights	0 - very unclear on weights
	13. Estimating resources and costs	Describes approaches used to estimate and value resource use, and any adjustments made to approximate to opportunity costs.	1 - Approach to costing and sources of data are clear, with quality discussed	0.5 - some information on costing missing, OR quality of data sources unclear / not discussed	0 - much information / discussion of costing data sources missing
	14. Currency, price date, and conversion	Reports the dates of the estimated resource quantities and unit costs, and methods for adjusting to the year/currency of analysis.	1 - currency dates and conversions clear	0.5 - minor issues of clarity	0 - currency dates OR conversions not reported
Modelling	**15. Choice of model**	Describes and give reasons for the specific type of decision-analytical model used.	1 - model is clearly described	0.5 - model description has some limitations	0 - model unclear
	16. Assumptions	Describes all structural or other assumptions underpinning the decision-analytical model.	1 - model assumptions clear, allowing reproducibility	0.5 - most model assumptions clear, but not reproducible	0 - model assumptions unclear
	17. Analytical methods	Describes all analytical methods supporting the evaluation (e.g. missing or censored data; population heterogeneity)	1 - explains important steps, e.g. how cost categories / types summed, dealt with missing data / outliers	0.5 - partially explained	0 - poorly explained
	18. Study parameters	Report the values, ranges, references (and, if used, probability distributions) for all parameters.	1 - values, ranges and references for input parameters reported	0.5 - some model inputs described, but not values, ranges, references	0 - input parameters poorly described
Results and	**19. Incremental costs and outcomes**	Reports mean values for the main categories of estimated costs and outcomes, as well as mean differences between comparator groups	1 - mean values for cost categories / outcomes reported	0.5 - costs/outcomes only reported in aggregate, not disaggregated by category	0 - very unclear reporting of costs / outcomes

Area	Reporting item (from Husereau et al., 2013)		Interpretation for scoring in this study (authors)		
	20. Characterising uncertainty	Describes the effects of sampling and/or model uncertainty, with impact of methodological assumptions (e.g. discount rate)	1 - parameter and structural uncertainty adequately characterised	0.5 - only parameter OR structural uncertainty adequately characterised	0 - limited characterisation of sources of uncertainty
	21. Characterising heterogeneity	If applicable, reports differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups.	1 - if sub-group analysis conducted, it is clear	0.5 - if sub-group analysis conducted, it is fairly clear	0 - if sub-group analysis conducted, it is unclear
Other	**22. Findings, limitations, generalisability**	Summarises key study findings and limitations, describing how results support the conclusions, and their generalisability.	1 - findings / conclusions clearly linked and limitations discussed	0.5 - conclusions unclear or not linked to findings, OR limitations not discussed	0 - conclusions and limitations unclear
	23. Source of funding	Describes how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis.	1 - funder noted	n/a	0 - funder not noted
	24. Conflicts of interest	Describes any potential for conflict of interest of study contributors in accordance with journal policy.	1 - COI adequately described	n/a	0 - COI not described

Figure 5: Histogram of CHEERS scores across n=53 studies



F. Original and inflation-adjusted results of cost-effectiveness studies

We followed guidance from Turner et al. (2019a) for adjusting costs with local inflation rates which have already been reported in US$. Our process was:

1. convert values into local currency for the respective country at the exchange rate in the study year
2. inflate to 2019 prices using local inflation
3. convert back into US$ using 2019 exchange rate.

Generic studies not in a specific country were adjusted using US inflation.

Table 11: Original and inflation-adjusted results of cost-effectiveness studies

Reference	ID	Year of analysis	Original base case results (US$ in year of analysis)	Inflation-adjusted base case results (2019 US$)
Varley (1998)	1020	1996	1,152 per DALY averted (ACER) 39,720 per death averted (ACER)	1,770 per DALY averted (ACER) 61,017 per death averted (ACER)
Shrestha (2006)	1036	2004	1,252 per DALY averted (ICER)	2,041 per DALY averted (ICER)
Cairncross (2006)	1037	2000	per DALY averted (ACER) (i) 94 (ii) 223 (iii) 47 (iv) 270 (v) 11	per DALY averted (ACER) (i) 135 (ii) 321 (iii) 68 (iv) 389 (v) 16
Clasen (2007)	1043	2002	per DALY averted (ACER) (i) 123 (ii) 53 (iii) 142 (iv) 61 (v) 472	per DALY averted (ACER) (i) 171 (ii) 74 (iii) 197 (iv) 85 (v) 655
Haller (2007)	1044	2000	per DALY averted (ACER) (i) 20 (ii) 571 (iii) 208 (iv) 687	per DALY averted (ACER) (i) 29 (ii) 822 (iii) 299 (iv) 989
Barungi (2011)	1071	2010	per DALY averted (ACER) (i) 25,551 (ii) 40,446	per DALY averted (ACER) (i) 22,457 (ii) 35,548
Murray (1998)	1183	1996	per DALY averted (i) stable: 433, refugee: 276, (ii) stable: 556, refugee: 313 (iii) stable: 327, refugee: 170	per DALY averted (i) stable: 665, refugee: 424, (ii) stable: 854, refugee: 481 (iii) stable: 502, refugee: 261
Jha (1998)	1022	1994	343 per life year saved (ACER)	467 per life year saved (ACER)
Larsen (2003)	1029	2000	per child death averted (ACER): (i) 1000 (ii) 2000	per child death averted (ACER): (i) 1,440 (ii) 3,072

Reference	ID	Year of analysis	Original base case results (US$ in year of analysis)	Inflation-adjusted base case results (2019 US$)
Meddings (2004)	1032	1999	per death averted (ACER): 1804 (provider perspective) 3436 (societal perspective)	per death averted (ACER): 4,862 (provider perspective) 9,260 (societal perspective)
Gunther (2013)	1092	2000	per life year saved (ACER): (i) 2,184 (ii) 664	per life year saved (ACER): (i) 3,145 (ii) 956
Spears (2013)	1099	2010	per infant death averted (ACER) - 2,817 per life year saved (ACER) - 35	per infant death averted (ACER): 2,761 per life year saved (ACER): 34
Bhalotra (2017)	1136	2015	per death averted (ACER): 65,981 per life year saved (ACER) - 1,100	per death averted (ACER): 66,175 per life year saved (ACER): 1,103
Rogers (2019)	1158	2016	Per additional child recovered (with ICERs relative to control): (i) ICER 24 , ACER 124 (ii) ICER 149, ACER 162 (iii) ICER 654, ACER 287	Per additional child recovered (with ICERs relative to control): (i) ICER 19 , ACER 100 (ii) ICER 120, ACER 131 (iii) ICER 529, ACER 232
Dupas (2020)	1179	2018	per diarrhoea case averted (ACER) (i) $4.90 (ii) $27.31 (iii) $20.02	per diarrhoea case averted (ACER) (i) $5.21 (ii) $29.02 (iii) $21.27

G. Uncertainty and heterogeneity

In our results presentation and discussion, we refer to different types of uncertainty and heterogeneity. For the most part we follow the Briggs et al. (2012) definitions (Table 12), as reported in Drummond et al. (2015).

Table 12: Types of uncertainty and heterogeneity, based on Briggs et al. (2012)

	Term	Concept	Analogous term in regression model
Types of uncertainty (Briggs et al., 2012)	Stochastic uncertainty (also variability)	Random variability in outcomes between identical individuals	Error term, cannot be reduced by collecting more data
	Parameter uncertainty	The uncertainty in estimation of the parameter of interest	Standard error of the estimate, can be reduced by collecting more data
	Structural uncertainty (also model uncertainty)	The assumptions inherent in the decision model	The form of the regression model (e.g., linear, log-linear, etc.)
Types of heterogeneity (used for this review)	Sub-group heterogeneity (simply "heterogeneity" in Briggs et al., 2012)	The variability between individuals that can be attributed to their characteristics	Beta coefficients (or the extent to which the dependent variable varies by individual characteristics)
	Location heterogeneity	The variability between possible settings in hypothetical studies, which can be attributed to the hypothesised characteristics of settings	n/a

However, in the literature studied in this review, there are several studies which assessed hypothetical interventions in generic settings using Monte Carlo simulations (Jeuland et al., 2009; Radin et al., 2020; Whittington et al., 2017, 2012, 2009). In their simulations, these analyses sought to reflect not only parameter uncertainty as defined in Table 12, but also aspects of structural uncertainty (e.g. varying discount rate with a uniform distribution). They also modelled heterogeneity within the simulation, by allowing input variables like "uptake" to vary based on hypothesised cross-setting differences rather than likely parameter uncertainty for a given setting.

For example, a CBA of CLTS in Ethiopia might model uptake using a point estimate and standard error from a previous study to represent parameter uncertainty. However, the "generic" studies cited above, aiming to represent settings across LMICs, have tended to set the range to cover all plausible levels of uptake in LMICs. We refer to this as "location heterogeneity". Such studies have therefore incorporated parameter uncertainty, structural uncertainty and location heterogeneity, all as part of the same Monte Carlo simulation. We refer to heterogeneity in the Briggs et al. (2012) sense as “sub-group heterogeneity”.

H. Types of benefit and valuation methods

Table 13 presents the types of benefit which have been measured and valued in WASH cost-benefit analysis (CBA) studies. It is separated into those which have frequently been valued and included in headline results, versus those which have not typically been quantified, or have been measured indirectly and not valued. A description is provided for each, as well as common valuation methods. Table 14, further below, presents the outcomes valued in included CBA studies, and the percentage of total benefits that each represents. Where studies compared multiple interventions, only one example is provided for illustrative purposes.

Table 13: Types of benefit valued in WASH economic evaluations

Impact area	Type of benefit	Description	Common valuation methods
Outcomes frequently valued and included in headline results			
Direct health costs	Mortality	The value of lost lives	Value of a statistical life (VSL), human capital approach (HCA), or 'years of life lost' (YLL) component of DALY (Robinson et al. 2019)
	Morbidity - cost of illness (COI) to the patient	Direct costs borne by patients/families (e.g. transport, fees, medicines)	Willingness to pay to avoid disease, financial expenditure or 'years lost due to disability' (YLD) component of DALY (Robinson & Hammitt, 2018)
	Morbidity - COI to the health system	Direct costs of treating patients (e.g. clinicians' time)	Opportunity cost of clinicians' time (Robinson & Hammitt 2018)
Indirect health costs	Lost time of patients	Indirect costs of being sick, e.g. lost wages, schooling or opportunity cost of time	Value of time, usually taken as a proportion of an appropriate wage (Whittington & Cook, 2018)
	Lost time of caregivers	Indirect costs of caring for the sick, e.g. lost wages or opportunity cost of time	

Impact area	Type of benefit	Description	Common valuation methods
	Future productivity of children	Impact of future earnings caused by impaired nutrition or missing days of school	Not commonly included. This typically involves an assumption of impact on future wages, e.g. that a z-score increase in height for age increases future wage by 8% (Dickinson et al., 2015), or that missing a small number of school days has a linear effect on future wages (Sklar, 2017).
Time savings	Time savings or convenience	The value of time saved when people start using water supplies closer to home, or using household toilets instead of OD sites.	Value of time, usually taken as a proportion of an appropriate wage (Whittington & Cook, 2018)
Other	Coping costs	Financial costs averted as a result of the intervention, such as expenditure on bottled water or water treatment, or avoided fees for public toilet use.	Financial cost of the coping mechanism. Any time savings would be captured above.
	Reuse	Fertiliser replacement through use of urine and compost from faeces.	The opportunity cost of purchasing fertiliser to achieve the equivalent nutrient value (Dasgupta et al., 2019)
Outcomes typically not quantified, or not included in headline results			
Quality of life	Users' quality of life directly related to service use	Factors commonly reported by users as benefits of a higher level of service. These include safety, dignity, convenience and social status in the case of both water supply and sanitation. Privacy and disgust are also important for sanitation.	Never yet directly valued , though sometimes asked about in surveys. For example, survey respondents are sometimes asked about their levels of satisfaction with different aspects of sanitation (Hutton et al., 2014) or their extent of agreement with statements about benefits of owning a toilet (Hutton et al., 2020). For water supply, the value of non-health "aesthetic" benefits has been estimated very indirectly, via assumptions about proportions of time savings benefits, with a downward correction to avoid double-counting of health benefits (Jeuland & Whittington, 2009; Whittington et al., 2017). This is argued to capture the non-health component of consumer surplus resulting from increased water use.
	Clean neighbourhoods	The value of living in a cleaner neighbourhood, beyond individual service use	Never valued, though sometimes asked about in surveys (Hutton et al., 2014)

Impact area	Type of benefit	Description	Common valuation methods
Property value	Property value	The value of a property being increased by the presence or improvement of a toilet	Rarely measured. Hutton (2020) includes it as a one-off benefit at the end of the toilet's useful life, by asking households what they think the increase in property value would be.
Tourism	Economic benefits of increased tourism	Increased tourism revenues when destinations (e.g. beaches, parks) are more attractive due to cleaner water or broader environments	Never valued, though sometimes asked about in surveys (Hutton et al., 2014)

Table 14: Benefits valued in included CBA studies and the percentage of total benefits

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Wyatt (1990)	1014	Infrastructure: diesel-powered deep boreholes	**Time savings (100%)**: avoided water collection time	n/a	Price premium for vended water (i.e. price difference divided by travel time to access water directly)
Hutton (2004)	1030	Infrastructure: (i) Water supply (ii) WSS (iv) WSS + HWT (v) piped water and sewerage	**Direct health benefits (13%)**: avoided mortality, avoided COI to patient, avoided COI to health system **Indirect health benefits (11%)**: gained work days, gained school days, gained caregiver days **Time savings (76%)**: avoided OD travel time (nb. for Africa D region, universal improved WSS scenario)	HCA	Work days gained / travel time: 100% of minimum wage

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Ignacio (2006)	1039	Infrastructure: (i) Basic ecosan (ii) Concrete ecosan (iii) Pit toilet (iv) Pour-flush to ST	**Willingness to pay (100%)**	n/a	n/a
Hutton (2007b)	1045	Infrastructure: (i) water (ii) sanitation (iii) WSS	**Direct health benefits (6%)**: avoided mortality, avoided COI to patient, avoided COI to health system **Indirect health benefits (18%)**: gained work days, gained school days, gained caregiver days **Time savings (77%)**: avoided OD travel time (nb. for Sub-Saharan Africa, WASH MDG average)	HCA	100% of GNI per capita
Hutton (2007a)	1046	Infrastructure: (i) water supply (ii) WSS (iii) WSS + HWT (iv) piped water and sewerage	**Direct health benefits (13%)**: avoided mortality, avoided COI to patient, avoided COI to health system **Indirect health benefits (5%)**: gained work days, gained school days, gained caregiver days **Time savings (82%)**: avoided OD travel time (nb. for Africa D region, WASH MDG scenario)	HCA	100% of GNI per capita
Reddy (2008)	1054	Infrastructure: Village-level UV water treatment plants	**Direct health benefits (41%)**: avoided COI to patient **Indirect health benefits (45%)**: gained work days **Coping costs (14%)**: avoided bottled water expenditure	n/a	n/a
Whittington (2009)	1055	(i) Infrastructure: borehole with hand pump (ii) CLTS: basic pit latrine (iii) Infrastructure: biosand filters	**Direct health benefits (34%)**: avoided mortality, avoided COI to patient **Time savings (49%)**: avoided water collection time **Aesthetic (8%):** non-health component of increased water use (nb. for borehole with handpump scenario. CLTS breakdown is 83%/17% between direct health and time savings. Biosand filter breakdown is 100% direct health)	VSL	30% of assumed wage

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Jeuland (2009)	1059	Infrastructure: (i) boreholes with hand pump: 3.2 (ii) biosand filters: 2.9	**Direct health benefits (50%)**: avoided mortality, avoided COI to patient **Time savings (30%)**: avoided water collection time **Aesthetic (21%):** non-health component of increased water use (nb. for borehole with handpump scenario. Biosand filter breakdown is 100% direct health)	VSL	30% of assumed wage
Rodriguez (2011)	1069	Infrastructure: (i) rural dry pit (ii) rural flush to pit (iii) rural flush to septic tank (iv) urban flush to pit (v) urban flush to septic tank (vi) urban flush to sewer	**Direct health benefits (40%)**: avoided mortality, avoided COI to patient **Indirect health benefits (11%)**: gained work days **Time savings (46%)**: avoided OD travel time **Coping (2%)**: avoided water collection & treatment (nb. for rural OD to basic scenario, some scenarios include reuse)	HCA	30% of GNI per capita for adults (15% for children)
Winara (2011)	1070	(i) rural flush to pit (ii) rural flush to septic tank (iii) urban flush to pit (iv) urban flush to septic tank (v) urban flush to sewer	**Direct health benefits**: avoided mortality, avoided COI to patient **Indirect health benefits**: gained work days **Time savings**: avoided OD travel time **Coping**: avoided water collection & treatment (nb. no table which permits calculation of % contributions, and authors' Table 29 is not clear on which scenario is presented)	HCA	30% of GNI per capita for adults (15% for children)

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Chuan (2012)	1081	(i) rural dry pit (ii) rural flush to septic tank (iii) urban dry pit (iv) urban flush to septic tank (v) urban flush to sewer	**Direct health benefits (58%)**: avoided mortality, avoided COI to patient **Indirect health benefits (9%)**: gained work days **Time savings (30%)**: avoided OD travel time **Coping (3%)**: avoided water collection & treatment (nb. for rural OD to basic scenario, some scenarios include reuse)	HCA	30% of GNI per capita for adults (15% for children)
Nguyen (2012)	1082	(i) rural flush to pit (ii) rural flush to septic tank (iii) urban flush to pit (iv) urban flush to septic tank (v) urban flush to sewer	**Direct health benefits (36%)**: avoided mortality, avoided COI to patient **Indirect health benefits (1%)**: gained work days **Time savings (28%)**: avoided OD travel time **Coping (36%)**: avoided water collection & treatment (nb. for urban OD to sewerage)	HCA	30% of GNI per capita for adults (15% for children)
Whittington (2012)	1084	(i) CLTS: basic pit latrine (ii) infrastructure: biosand filters	**Direct health benefits (~80%)**: avoided mortality, avoided COI to patient **Time savings (~20%)**: avoided water collection time (nb. for CLTS. The biosand filter intervention benefits are 100% direct health)	VSL	30% of assumed wage
Fahimuddin (2012)	1085	Infrastructure: public taps and private connections	**Time savings (100%)**: avoided water collection time	n/a	52% of average wage (based on time use survey, whereby waged work weighted 100%, housework 50% and leisure 25%).

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Heng (2012)	1087	Infrastructure: (i) rural dry pit (ii) rural flush to pit (iii) urban flush to pit (iv) urban flush to sewer	**Direct health benefits (43%)**: avoided mortality, avoided COI to patient **Indirect health benefits (10%)**: gained work days **Time savings (23%)**: avoided OD travel time **Coping (24%)**: avoided water collection & treatment (nb. for rural OD to pit latrine)	HCA	30% of GNI per capita for adults (15% for children)
Hutton (2013)	1089	Infrastructure: (i) Water supply (mix - see table 2) (ii) Sanitation (mix - see table 2)	**Direct health benefits (14%)**: avoided mortality, avoided COI to patient **Indirect health benefits (5%)**: gained work days **Time savings (81%)**: avoided OD travel time (nb. for universal sanitation, water breakdown is 26% / 6% / 71% respectively)	HCA	30% of GDP per capita for adults (15% for children)
Rodriguez (2013)	1097	Infrastructure: (i) rural dry pit (ii) rural flush to pit (iii) rural flush to septic tank (iv) urban dry pit (v) urban flush to pit	**Direct health benefits (35%)**: avoided mortality, avoided COI to patient **Indirect health benefits (4%)**: gained work days **Time savings (57%)**: avoided OD travel time **Coping (4%)**: avoided water collection & treatment (nb. for rural OD to basic)	HCA	30% of GNI per capita for adults (15% for children)
Cronin (2014)	1105	Infrastructure - pour-flush to off-set pit	**Direct health benefits (77%)**: avoided mortality, avoided COI to patient **Indirect health benefits (9%)**: gained work days **Time savings (20%)**: avoided OD travel time	unclear (appears HCA)	100% of state GDP per capita (separate values for rural and urban)
Boije (2014)	1108	Infrastructure: (i) atmospheric water generator (ii) atmospheric heat-pump (iii) solar disinfection	**Direct health benefits (43%)**: avoided mortality and COI **Indirect health benefits (27%)**: gained work days & school days **Time savings (29%)**: avoided water collection time	VSL	100% of mean wage (from survey data)

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Dickinson (2015)	1109	CLTS with subsidy - typically achieved pour-flush to off-set pit	**Direct health benefits (40%)**: avoided mortality and COI **Time savings (16%)**: avoided water collection time **Nutrition (16%)**: improved height-for-age increased future wages	VSL	30% of mean wage (from survey data)
Hutton (2015)	1112	Infrastructure: (i) rural water (mix - see table 2) (ii) rural sanitation (mix - see table 2)	**Direct health benefits (18%)**: avoided mortality, avoided COI to patient **Indirect health benefits (14%)**: gained work days **Time savings (68%)**: avoided water collection time (nb. for urban water - the breakdown for urban sanitation is 14% / 9% / 77%)	HCA	30% of GDP per capita (15% for school-age children and caregivers)
Larsen (2016)	1122	Infrastructure: (i) deep tubewell (ii) pond sand filter (iii) rainwater harvesting (iv) improved pit latrine (from unimproved) (v) improved pit latrine (from shared)	Unclear methods, but appears **Direct health benefits (100%)**: combining avoided mortality from arsenic exposure and diarrhoea	HCA	50% of national mean wage (ILO data)
Meeks (2017)	1128	Infrastructure: public taps	**Time savings (100%)**: avoided water collection time	n/a	100% of wage (estimated from calculations of returns to labour)
Whittington (2017)	1131	(i) Infrastructure: borehole with handpump (ii) Infrastructure: borehole with handpump & biosand filter (iii) CLTS: improved pit latrine	**Direct health benefits (52%)**: avoided mortality and morbidity via DALYs **Time savings (46%)**: avoided water collection time **Aesthetic (2%):** non-health component of increased water use (nb. for borehole with handpump including biosand filter scenario.	DALY valued at 3 x GDP per capita	50% of assumed unskilled wage

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
			CLTS breakdown is 67%/33% between direct health and time savings.)		
Sklar (2017)	1132	(i) Infrastructure: pour-flush to septic tank (ii) Service lease: container-based sanitation	**Direct health benefits (71%)**: avoided mortality and morbidity via DALYs **Time savings and productivity (26%)**: avoided OD travel time or public toilet queuing time, and work days lost (not reported separately) **Education (2%):** future earning increased when school not missed nb. benefits are assumed identical for the two options	DALY valued at 3 x GDP per capita	50% of GDP per capita
Larsen (2018)	1140	Infrastructure: (i) Urban improved water supply (ii) Rural improved water supply (iii) Urban improved sanitation (iv) Rural improved sanitation Promotion: (v) Use of existing toilets (vi) HWT (filter)	**Direct health benefits (54%)**: avoided mortality, avoided COI **Indirect health benefits (5%)**: gained work days **Time savings (41%)**: avoided OD travel time (nb. for subsidised rural sanitation. Breakdown for sanitation promotion is 32% / 4% / 64%, while for POU water treatment it is 89% / 11% for direct and indirect health respectively)	VSL	Sick days & travel time: 50% of wage (estimated from GDP per capita, labour force participation, and labour income share of GDP)
Cha (2018)	1143	Infrastructure: (i) new borehole, (ii) rehabilitated borehole	**Direct health benefits (19%)**: avoided mortality, avoided COI to patient, avoided COI to health system **Indirect health benefits (13%)**: gained work days / caregiver days **Time savings (68%)**: avoided water collection time (nb. for new borehole with handpump scenario)	HCA	100% of mean caregiver wage (from survey data)

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Burt (2018)	1145	Infrastructure - upgrade existing water network to continuous water supply	**Avoided coping expenditures (unclear)**: avoided household investment and recurrent costs of coping with IWS **Time savings (unclear)**: avoided time spent waiting for water to be switched on	n/a	50% of wage (from survey data?)
Larsen (2018)	1149	Infrastructure: (i) Urban improved water supply (ii) Rural improved water supply (iii) Urban improved sanitation (iv) Rural improved sanitation Promotion: (v) Use of existing toilets (vi) HWT (filter)	as for Larsen 2018a above	VSL	Sick days & travel time: 50% of wage (?)
Dasgupta (2019)	1162	Infrastructure: (i) eco-san (ii) twin-pit pour flush toilet	**Direct health benefits (22%)**: avoided COI **Indirect health benefits (36%)**: gained work days **Time savings (33%)**: "convenience" benefits **Reuse (9%)**: fertiliser replacement through use of urine and compost (nb. for the ecosan option, pit latrine breakdown was 14% / 22% / 65% for the first 3 categories)	n/a	100% of mean wage (from survey data)
Das (2019)	1167	Infrastructure - pond-sand filters	**Direct health benefits (23%)**: avoided COI **Indirect health benefits (34%)**: gained work days **Coping cost (33%)**: avoided expenditure on water treatment / bottled water	n/a	Sick days: 100% of mean wage (from survey data)

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Hutton (2017)	1173	Promotion campaign with ex post subsidy - flush to offset pit / septic tank.	**Direct health benefits (51%)**: avoided mortality and COI **Indirect health benefits (4%)**: gained work days **Time savings (45%)**: avoided access time	VSL	100% of reported wage, or rural unskilled wage if unwaged (50% for 5-14, 0% for <5) caregivers time if <5 sick
Radin (2020)	1175	CLTS: improved pit latrine	**Direct health benefits (67%)**: avoided mortality and COI **Time savings (33%)**: avoided OD travel time (nb. medium uptake scenario without externality)	VSL	50 % of assumed unskilled wage for adults (25% for 5-14, 0% for <5)
Dwumfour-Asare (2020)	1178	Infrastructure: (i) full subsidy with bylaw enforcement and 75% uptake (ii) subsidy only with 10% uptake	**Direct health benefits (13%)**: avoided mortality and COI **Time savings (64%)**: avoided public toilet travel/queuing time **Coping cost (23%)**: avoided fees for public toilet use (nb. for subsidy with enforcement scenario, under medium uptake)	VSLY	50% of average urban wage for adults (25% for 5-14, 0% for <5)
Reeser (1988)	1180	Diesel-powered deep boreholes	**Time savings: 100%**	n/a	47% of minimum wage (100% * (1-unemployment rate) * shadow price rate for unskilled labour)
Abelin (1997)	1181	Infrastructure : handpumps	**Time savings: 100%**	n/a	60% of unskilled wage
Eklund (1991)	1184	Infrastructure: gravity scheme with tapstands	**Time savings: 100%**	n/a	100% of wage (weighted by seasonal availability of farm work)

Reference	ID	Intervention and technology		Methods for value of life	Methods for value of time
Cha (2020)	1185	CLTS: Improved pit latrine	**Direct health benefits (64%)**: avoided mortality, avoided COI to patient and health system **Indirect health benefits (7%)**: gained work days **Time savings (29%)**: avoided access time to shared latrines / OD	VSL	50% of mean wage, from survey data (25% for 5-14, 0% for <5)

Appendix B: Supplementary materials for Chapter 5 (qualitative paper)

A. Reporting against COREQ checklist
B. Further information about the intervention
C. IDI topic guide
D. Additional findings
E. Original Portuguese of quotations

A. Reporting against COREQ checklist

No.	Item	Guide questions/description	Page no.	Notes
Domain 1: Research team and reflexivity				
Personal Characteristics				
1.	Interviewer / facilitator	Which author/s conducted the interview or focus group?	See published version (SPV)	n/a – we worked with contracted interviewers with no interest in academic publication, due to Changana language requirements.
2.	Credentials	What were the researcher's credentials? *E.g. PhD, MD*	n/a	Two interviewers had an undergraduate degree, and the other two a secondary education.
3.	Occupation	What was their occupation at the time of the study?	n/a	Interviewers worked part-time alongside other various employment
4.	Gender	Was the researcher male or female?	SPV	Two male, two female
5.	Experience and training	What experience or training did the researcher have?	n/a	All had previous experience in social research, and two had experience facilitating focus groups. The week's training is described in the main body.
Relationship with participants				
6.	Relationship established	Was a relationship established prior to study commencement?	n/a	Individual interviewers had no established relationship with participants before the study, but [initials]'s firm have been undertaking MapSan data collection activities in these bairros since 2015.

No.	Item	Guide questions/description	Page no.	Notes
7.	Participant knowledge of the interviewer	What did the participants know about the researcher? e.*g. personal goals, reasons for doing the research*	n/a	We see no reason why participants would expect interviewers to have any personal goals related to the research. Reasons for the research were in the participant information sheet explained at the beginning of each data collection event.
8.	Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. *Bias, assumptions, reasons and interests in the research topic*	n/a	None. They introduced themselves as employees of [company]. Their clothes or manner of speaking may have led participants to judge their social status.
Domain 2: study design				
Theoretical framework				
9.	Methodological orientation and Theory	What methodological orientation was stated to underpin the study? *e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis*	SPV	interpretivist position applying framework analysis, and initial coding drawing on grounded theory
Participant selection				
10.	Sampling	How were participants selected? *e.g. purposive, convenience, consecutive, snowball*	SPV	full detail at referenced location
11.	Method of approach	How were participants approached? e.*g. face-to-face, telephone, mail, email*	SPV	full detail at referenced location
12.	Sample size	How many participants were in the study?	SPV	See referenced location. nb. only one participant was sampled per compound.

No.	Item	Guide questions/description	Page no.	Notes
13.	Non-participation	How many people refused to participate or dropped out? Reasons?	n/a	Refusal to participate was 5% for IDIs, which were arranged to suit the respondent, and 15% for FGDs, which took place on Saturday mornings.
Setting				
14.	Setting of data collection	Where was the data collected? e.*g. home, clinic, workplace*	n/a	IDIs took place in participants' home compounds during the week. FGDs took place on Saturdays in community buildings, such as nurseries and NGO offices.
15.	Presence of non-participants	Was anyone else present besides the participants and researchers?	n/a	For IDIs, non-participants (particularly children) were occasionally within earshot. FGDs took place without any non-participants present
16.	Description of sample	What are the important characteristics of the sample? *e.g. demographic data, date*	Tab. 1	full detail at referenced location
Data collection				
17.	Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	SPV	full detail at referenced location and guide in supplementary materials C
18.	Repeat interviews	Were repeat interviews carried out? If yes, how many?	n/a	no, but last 2 x FGDs reconvened previous participants
19.	Audio/visual recording	Did the research use audio or visual recording to collect the data?	SPV	Audio-recording. Any names or other identifiers included in audio recordings were redacted during transcription. All audio recordings were permanently deleted

No.	Item	Guide questions/description	Page no.	Notes
20.	Field notes	Were field notes made during and/or after the interview or focus group?	n/a	Interviewers took limited notes in IDIs, in order to focus on the flow of the discussion.
21.	Duration	What was the duration of the interviews or focus group?	SPV	full detail at referenced location
22.	Data saturation	Was data saturation discussed?	SPV	full detail at referenced location
23.	Transcripts returned	Were transcripts returned to participants for comment and/or correction?	n/a	no
Domain 3: analysis and findings				
Data analysis				
24.	Number of data coders	How many data coders coded the data?	SPV	One coder. In addition, [initials] speaks Portuguese to a level sufficient for understanding original transcripts as well as translations. [initials] and [initials] discussed, with interviewers, the meanings of terms used by participants, to establish whether the way they were interpreted in English was the same as in Portuguese.
25.	Description of the coding tree	Did authors provide a description of the coding tree?	Fig. 2	The conceptual model comprises the final core codes
26.	Derivation of themes	Were themes identified in advance or derived from the data?	SPV	derived from the data
27.	Software	What software, if applicable, was used to manage the data?	SPV	nVivo 12

No.	Item	Guide questions/description	Page no.	Notes
28.	Participant checking	Did participants provide feedback on the findings?	n/a	During the participant checking process in the last two FGDs, no substantial concerns or proposals were raised.
Reporting				
29.	Quotations presented	Were participant quotations presented to illustrate the themes / findings? Was each quotation identified? e.*g. participant number*	SPV	full detail at referenced location
30.	Data and findings consistent	Was there consistency between the data presented and the findings?	SPV	full detail at referenced location
31.	Clarity of major themes	Were major themes clearly presented in the findings?	Fig. 2	full detail at referenced location
32.	Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Fig. 2	full detail at referenced location

B. Types of toilets

The two types of toilets delivered as the MapSan trial intervention were, firstly, a single 'shared toilet' to be used by a minimum of 15 people, at 85% subsidy. The second is a community sanitation block to be used by a minimum of 21 people, at 90% subsidy. All households had on-plot piped water supply at the time of intervention, and CSBs have their own water tank on the roof. The NGO has been implementing the same intervention subsequent to MapSan in the same neighbourhoods. They have been iteratively improving variants of this intervention in low-income areas of Maputo since 2009. Photographs with typical examples of each of the three toilet types are below.

Pit latrines

1. Pit latrine with tyre and wood for squatting	2. Pit latrine with concrete slab
	
3. Fabric door providing limited privacy	4. No door and adjacent greywater pit
	

Shared toilets and community sanitation blocks

Exterior	
1. Shared toilet (ST)	2. Community sanitation block (CSB)
	
Interior (varied between CSB / ST depending on design)	
3. Squat pan	4. Seat pan
	

Further information about the intervention

We provide more information in the table below about the intervention, in the format of the TIDieR checklist (Hoffmann et al., 2014). More information about different aspect of the intervention is provided across various publications (Bick et al., 2020; Knee et al., 2020; Mattson, 2016).

Intervention description using TIDieR checklist

	Item	Notes
1	Provide the name or a phrase that describes the intervention.	Subsidised pour-flush toilets shared by multiple households
2	Describe any rationale, theory, or goal of the elements essential to the intervention.	In this setting there is limited space and willingness or ability to pay for private toilets, and households already use low-quality shared pit latrines.
3	Materials: Describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or in training of intervention providers. Provide information on where the materials can be accessed (e.g. online appendix, URL).	The intervention provided two types of toilet facility (photos above), alongside education on their use and maintenance. There were also two hygiene promotion visits after completion of construction, carried out by paid staff who received 2 days of training. These focused on contamination routes, good personal hygiene practice, and handwashing with soap. More information is provided elsewhere (Bick et al., 2020; Knee et al., 2020; Mattson, 2016).
4	Procedures: Describe each of the procedures, activities, and/or processes used in the intervention, including any enabling or support activities.	Key procedures included: 1. Community engagement and site identification – undertaken by eight contracted community-based organisations (CBOs), e.g. assessment of demand for better toilets and localised environmental issues affecting site selection (e.g. water table) 2. Site selection and preparation – site selection undertaken by WSUP in discussion with CBOs, and site preparation (e.g. emptying of old latrine pits) undertaken by contracted firms. 3. Toilet construction – undertaken by contracted construction firms 4. Education on use, maintenance and hygiene – undertaken by contracted 'sanitation activists'
5	For each category of intervention provider (e.g. psychologist, nursing assistant), describe their expertise, background and any specific training given.	Main stakeholders in delivery included: 1. **Water and Sanitation for the Urban Poor** (international NGO) – overall lead on intervention delivery. Team included engineers and community engagement specialists. 2. **Various community-based organisations** – sub-contractor facilitating community engagement. 48 people trained. Teams included facilitators from the local area of the intervention. 3. **Various construction firms** – Sub-contractors building the toilet infrastructure. They were predominantly small local firms. 4. **Sanitation activists** – Sub-contractors educating toilet users and promoting hygiene. 55 people trained. 5. **Municipality and World Bank** – oversight and approvals. Team included engineers.
6	Describe the modes of delivery (e.g. face-to-face or by some other mechanism, such as internet or telephone) of the intervention and whether it was provided individually or in a group.	All engagement was face-to-face. As this was shared sanitation, any site visits were made to compound members jointly, rather than individually.

7	Describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features.	Setting described fully in manuscript main body.
8	Describe the number of times the intervention was delivered and over what period of time including the number of sessions, their schedule, and their duration, intensity or dose.	All aspects of the intervention delivered only once, except for two hygiene promotion visits.
9	If the intervention was planned to be personalised, titrated or adapted, then describe what, why, when, and how.	n/a
10	If the intervention was modified during the course of the study, describe the changes (what, why, when, and how).	n/a
11	Planned: If intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them.	n/a
12	Actual: If intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned.	Fidelity was assessed by Bick et al. (2020)

C. IDI topic guide

Part 1. What is a good life? (15 mins)

We would like to talk about things that are important for a good life, such as:

- things you can buy and touch (for example, food)
- things you cannot buy (for example, family)
- things you can feel but cannot touch (for example, happiness)

What do you think is important for a good life?

Think about people who live in this neighbourhood who have a good life.
List ten things that make these people have a good life.

[probe – remind about things you cannot buy or touch]

Now think about people who live in this neighbourhood who have a bad life:

[probe – don't just think about the opposites of things for a good life]

Part 2. Contribution of sanitation to a good life (20-40 mins)

Let's talk about toilets and sanitation.
What types of toilets do people use in this neighbourhood?

[note – keep this short, just to establish that there are many types]

Earlier you made a list of things important for a good life.
Now let's talk about how good sanitation and good toilets contribute to each of these things. They can also be examples of bad toilets or sanitation.
Think about all things you use the toilet for.

Let's talk about [*QoL card X*] that you mentioned.

[QoL cards include: Enough food, Education for children, Security, Housing / shelter, Happy family and children, Partner, Friends / neighbours, Physical health, Mental health / peace of mind, Clean environment, Water and Sanitation, Good work / job, Enough money to live, Freedom / independence, Political voice, Happiness, Being respected]

How do you think that using good or bad toilets affect [*QoL card X*]?
It can be positive, negative, or maybe the answer is "not at all".

Think about how the effect can be different for people of different ages and sexes.

Now let's consider [*QoL card Y*]

[continue until all cards have been addressed, or until you run out of time]

[If the conversation is not flowing, pick a random QoL card which they didn't mention, emphasising that these are what others said was important. Alternatively, ask questions such as "are you satisfied with the type of bathroom you are using now?" or "If you improved your toilet, how would your quality of life change?"]

Can you think of other ways that good toilets contribute to a good life, that are not on a card?

Part 3. Pile-sorting component (5-10 mins)

These are things that other people said were important about sanitation.
You have already mentioned some of them.
Let's classify them according to their importance to live a good life.

First, select the five things that are the most important for a good life.

[probe - why did you make that choice? What makes these things more important?]

Now, select the five things which are next most important.

[probe - why did you make that choice? What makes these things more important?]

Close

An example of a pile-sorting card is provided below – the label reads “Not smelling faeces”. The 15 cards comprise the bar labels in Figure 4.

Não cheirando fezes

D. Additional findings

Matching concepts from the literature onto attributes

The identified attributes align to a great extent with the 15 concepts used for the pile-sorting cards identified from the literature (**Error! Reference source not found.**). As noted in relation to Figure 4, concepts which do not fit under any single attribute scored lower in pile-sorting than the highest-ranked concepts of our five core attributes, and were in the bottom 40% of cards overall. Notes on these pile-sorting concepts are provided in **Error! Reference source not found.**.

How pile-sorting concepts from the literature match onto identified attributes

Identified attribute	Concept on pile-sorting card from literature
Health	Fewer diseases
Disgust	Seeing faeces less
	Better personal cleanliness
	Fewer flies
	Smelling faeces less
Shame	More dignity
	More pride
	Less shame when visitors come
Safety	More safety for women and girls
Privacy	More privacy

Notes on pile-sorting concepts which do not map onto a specific attribute

Concept on pile-sorting card from literature	Notes on why it does not map under any single attribute
More comfort	Depending on the way the concept is used, the associated feeling or experience could conceivably be included under any of the five attributes. For example: • disgust, e.g. feeling more comfortable or at ease when not seeing, smelling or touching disgusting things • safety, e.g. feeling more comfortable or at ease when not worried about having an accident or being assaulted
Cleaner environment	As distinct from disgust related to specifically using the toilet, or its direct local consequences (e.g. overflowing latrines), the concept of a clean environment is much broader. When mentioned by participants, it was sometimes framed in terms of solid waste. In terms of sanitation, it relates more to environmental conditions arising from excreta in floodwater or illegal dumping of faecal sludge, rather than sanitation

	behaviours within the household. Therefore, the concept of a clean environment would either be included under disgust or not be a part of SanQoL as defined in this paper.
More convenience	See discussion section of manuscript
Easier for people with restricted mobility	In capability terms, restricted mobility is a conversion factor. Therefore, depending on the way the concept is used, the associated feeling or experience could conceivably be included under any of the five attributes. For example: • disgust, e.g. being less able to avoid seeing, smelling or touching disgusting things due to restricted mobility • shame, e.g. being ashamed or feeling less dignified as a result of being less able to carry out sanitation practices than one would like
Reduced conflict with neighbours	Depending on the way the concept is used, the associated feeling or experience could be included under: • shame, e.g. if conflict about sanitation arises from a behaviour or its consequence considered shaming • disgust, e.g. if conflict arises because the toilet is disgusting neighbours within or outside the compound

Triads methods

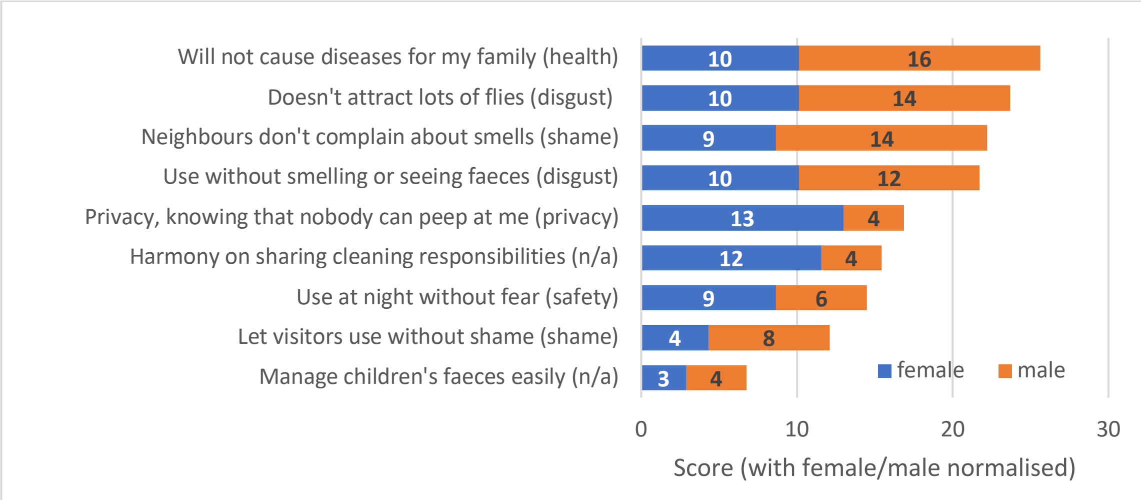
Participants were shown a set of nine A4 cards, each with an attribute of sanitation-related QoL identified as most important in emerging analysis of transcripts. A sub-set of three cards was then read out by an interviewer and placed on the floor for all to see. Participants were asked to choose the attribute they considered the most important of the three, and to tick the relevant box using a sheet and pen provided. The vast majority of participants were literate, but interviewers supported the few who struggled. This was repeated 15 times with different combinations of attributes. For the analysis, concepts were allocated to three groups based on the frequency with which they were preferred. The same scoring as IDIs was then applied (first group: 3 points, second: 1 point, third: 0 points).

Triads results

The triad exercise in the final two FGDs generated 195 observations. Results were broadly similar to the pile-sorting, except concepts linked to shame scored higher. "Managing children's faeces easily" was included as a sense-check as a concept that was being mentioned in emerging analysis but not that frequently. As expected, it scored low. "Harmony on sharing cleaning responsibilities" was more important in transcripts and scored especially high amongst women (who were more likely to carry out cleaning).

Triadic comparison of attributes during 2 x FGDs (n=13 respondents)

nb. labels (without bracketed part) were written on A4 cartoon cards in Portuguese and read out by interviewers. n/a = not applicable.



E. Original Portuguese of quotations

Original Portuguese of quotations used in text.

English	Portuguese
Health	**Saúde**
“You feel under pressure when the bathroom is dirty, and you don’t feel at ease.” Female FGD, 18-24 (FGF01)	"Sentes um aperto quando a casa de banho fica suja e não se sentes a vontade"
“It is difficult for us to control what children do. An adult knows they shouldn’t touch something, or they'll catch germs, but a child doesn't know.” Male IDI, 36 (EAGJ04)	"é difícil nós controlarmos as crianças, o sitio onde um adulto sabe que não posso pegar aqui senão vou ter micróbios a criança não sabe"
“When it rains the faeces in the pit rise up, then we get diseases because cholera comes from there” Female IDI, 71 (EANC04)	"Quando chove as fezes sobem e de repente te deparas com aquelas larvas nesse momento traz doenças dentro de casa como cólera"
“Having a good toilet contributes positively to all these aspects, mental health, wellbeing for the soul, and general health as a whole.” Male IDI, 64 (EAGJ03)	“ter uma boa casa de banho contribui positivamente para todos esses aspetos que me referia, saúde mental, contribui para o bem estar para a alma, contribui para a saúde em geral no seu todo”
“Your neighbours will know the origin of the smell and will start to talk about it, and you can’t feel relaxed.” Female IDI, 76 (EAET04)	“vão saber a origem do cheiro e os vizinhos irão começar a falar e isso não te pode deixar tranquila”
Disgust	**Nojo**
"It is something so horrible to see other people's faeces." Male IDI, 19 (EAGJ02)	"é algo tão horrivel estar a ver fezes de outras pessoas"
“You cannot eat because you lose your appetite … due to the smell. You don’t even feel free to come out of your house because it smells bad out there.” Female FGD, 25-59 (FGF02)	“Mesmo comer você não pode, porque perde apetite … devido ao cheiro. Nem tem vontade de sair de dentro de casa porque fora cheira mal.”
“[In the toilet] I would feel like I am in the kitchen. With no bad smell, it seems like you could even drink tea in there, without realising you are in a toilet.” Female IDI, 71 (EANC04)	"Parece que estou a entrar na cozinha, não cheira nem nada até parece que vais entrar na casa de banho para tomar chá sem se aperceber que ali é uma casa de banho"
When the house is clean but the toilet is not, this is undignified.” Female FGD, 60+ (FGF03)	“quando a casa fica limpa e a casa de banho não, não traz dignidade.”

Shame	Vergonha
"A person's toilet becomes the mirror of that person." Male IDI, 64 (EAGJ03)	"A casa de banho da pessoa torna-se o espelho da pessoa"
"Everyone will refer to you according to the state of your toilet, saying 'it's there at her house that the toilet smells' … nobody respects you." Female IDI, 27 (EAET05)	"Toda gente irá referenciar te em função do estado da sua casa de banho, dizendo é ali em casa da dona *[name redacted]* que cheira casa de banho … ninguém te respeita."
"People who go down my road smell the stench from my toilet. Then when they later pass me on the street they will look at me in a different way." Male IDI, 19 (EAGJ02)	"Quem passa da minha rua e ao lado da minha casa sentisse o meu mau cheiro que parte da casa de banho da minha casa, essa pessoa, no dia que for a cruzar-se comigo na rua epah, me repararia de uma outra forma"
"When I get visitors, I can let the person use the toilet without fear. I think this makes people look at me differently, with respect." Male IDI, 28 (EAJP05)	"Recebo visitas, posso muito bem deixar que a pessoa use a casa de banho sem receio, então acho que isso acaba fazendo com que as pessoas olhem-me de outra forma, com respeito"
"If a visitor asks to go to the toilet and sees it in good condition, they'll say 'wow, that lady's house is hygienic'" Female FGD, 60+ (FGF03)	"Chegar alguém e pedir para ir a casa de banho, e quando chega lá, percebe que está em condições. Na saída dele vai dizer 'hiii em casa de fulano há higiene'"
Safety	**Segurança**
"There are people who are raped while they use these toilets. … there are times we even have to defecate in a bucket because we fear bandits." Female IDI, 27 (EAET05)	"Existem pessoas que são violadas por causa de usarem estas casas de banho, é por isso que a noite usamos baldinhos para o efeito, temendo violência"
"that toilet built from car tyres is a hazard – when it rains it could come crashing down at any moment." Male FGD, 25-59 (FGM02)	"Aquele pneu é um risco a qual, quando chove a qualquer momento aquilo pode desabar."
"I'm afraid to use it at night because I wouldn't know which way to enter, where to tread inside, and I would be afraid of falling into the hole." Female IDI, 71 (EANC04)	"Não posso ter coragem porque ao sair de noite não saberia como entrar, onde pisar dentro da casa de banho e tenho medo de entrar ali no buraco"
Privacy	**Privacidade**

“While you walk to work, … you might see a woman with just a bit of capulana [fabric], when she is naked taking a bath.” Male IDI, 28 (EAJP05)	"A gente quando vai ao trabalho, ... dá para perceber uma mulher têm um bocado de capulana, quando é que está despida no banho"
“When a bathroom is not secure you do not feel free to use it, because at any moment an individual can enter.” Female IDI, 76, (EAET04)	"Quando uma casa de banho não oferece segurança voce não esta livre de utiliza-la porque a qualquer momento poderá entrar um individuo dentro"
“You cannot imagine the gymnastics I do when I have my period. I do not feel relaxed because I do not know if I'm being watched.” Female IDI, 27 (EAET05)	"Não imaginas a ginástica que eu faço quando estou nos dias da minha menstruação, não me sinto a vontade porque não sei se estou sendo observada"

Appendix C: Supplementary materials for Chapter 6 (measure development paper)

A. More information on setting

B. SanQoL items in Portuguese

C. Robustness checks on ranking and hypothesis testing

D. Visual analogue scale

E. Reporting against ISOQOL criteria

A. More information on setting

Pre-intervention, all households used a pit latrine shared with other households. The majority of these were 'traditional latrines' with a soil floor, are called *casa de banho* (bathroom) as the space is also used for bathing. We use 'toilet' to refer to all locations where defecation or bathing took place, for consistency, but we explicitly recognise that these facilities are used for other purposes, such as menstrual hygiene management. The intervention was delivered during 2015-16 by Water and Sanitation for the Urban Poor (WSUP), an international non-government organisation. MapSan intervention compounds were provided with a highly-subsidised pour-flush toilet discharging to a septic tank, of two design types depending on the number of users. Control compounds continued to use shared pit latrines.

The two types of toilets delivered as the MapSan trial intervention were, firstly, a single 'shared toilet' to be used by a minimum of 15 people, at 85% subsidy. The second is a community sanitation block (CSB) to be used by a minimum of 21 people, at 90% subsidy. All households had on-plot piped water supply at the time of intervention, and CSBs have their own water tank on the roof.

Pit latrines

1. Pit latrine with tyre and wood for squatting	2. Pit latrine with concrete slab
	

3. Fabric door providing limited privacy	4. No door and adjacent greywater pit
	

Shared toilets and community sanitation blocks

Exterior	
1. Shared toilet (ST)	2. Community sanitation block (CSB)
	
Interior (varied between CSB / ST depending on design)	
3. Squat pan	4. Seat pan
	

B. SanQoL items in Portuguese

Table 15: SanQoL items in English and Portuguese

Attribute		Questionnaire item	Responses
1	Disgust ***Nojo***	Can you use the toilet without feeling disgusted? ***Pode usar a casa de banho sem se sentir nojo?***	1 - Always 2 - Sometimes 3 - Rarely 4 - Never ***1 - Sempre*** ***2 - Às vezes*** ***3 - Raramente*** ***4 - Nunca***
2	Health ***Saúde***	Can you use the toilet without worrying that it spreads diseases? ***Pode usar a casa de banho sem se preocupar que espalhe doenças?***	
3	Privacy ***Privaçidade***	Can you use the toilet in private, without being seen? **Pode usar a casa de banho com privacidade, sem ser visto/a?**	
4	Shame ***Vergonha***	Can you use the toilet without feeling ashamed for any reason? ***Pode usar a casa de banho sem sentir vergonha por qualquer motivo?***	
5	Safety ***Segurança***	Are you able to feel safe while using the toilet? ***É capaz de se sentir seguro/a ao usar esta casa de banho?***	

C. Robustness checks on ranking and hypothesis testing

Table 16: Ranking robustness check using GLMM (ranks modelled as continuous)

		Disgust	Health	Shame	Safety	Privacy
Female	Coeff.	-0.055	0.008	0.137	0.058	-0.146
	p-value	0.643	0.929	0.177	0.622	0.165
Aged 60+	Coeff.	0.099	-0.262**	0.047	0.265	-0.132
	p-value	0.579	0.037	0.745	0.108	0.441
Treatment	Coeff.	-0.204	0.000	0.101	0.198	-0.085
	p-value	0.117	0.997	0.401	0.136	0.492

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

Table 17: Hypothesis robustness check using mixed effects ordered logit

Variable	**Type of variable**	**P-values on coefficients in mixed effects ordered logit regression on attribute scores**				
		Disgust	**Health**	**Shame**	**Safety**	**Privacy**
User characteristics						
Female	Binary	0.895	0.372	0.495	***<0.001****	0.772
Aged 60+	Binary	0.851	0.683	0.851	0.650	0.646
Wealth index	Continuous	***0.046***	0.269	0.426	0.305	0.975
Toilet characteristics						
Toilet floor material	Binary	0.489	***0.007****	***0.011***	***0.001****	***0.003****
Toilet wall material	Binary	0.107	0.129	0.727	***0.024***	***0.013***
Toilet inside lock	Binary	***0.020***	0.705	***0.006****	***0.003****	***<0.001****
Enumerator smells faeces	Binary	***0.005****	***<0.001****	***<0.001****	0.104	0.063*

Table 18: Results in agreement with presence hypotheses

Variable	Type of variable	Test results in line with hypothesis?				
		Disgust	Health	Shame	Safety	Privacy
User characteristics						
Women	Binary				yes	no
Aged 60+	Binary					
Wealth index	Continuous					
Toilet characteristics						
Toilet floor material (high/low quality)	Binary	no	yes	yes	yes	
Toilet wall material (high/low quality)	Binary			no	yes	yes
Toilet inside lock	Binary				yes	yes
Enumerator smells faeces	Binary	yes	yes	yes		

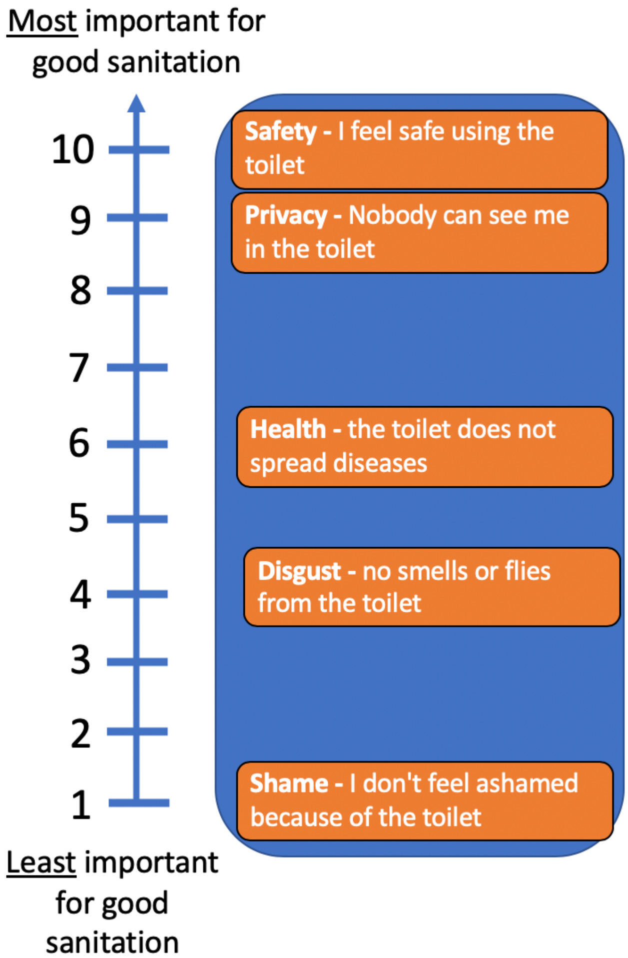
Figure 6: visualisation of mean rank data for overall sample

D. Visual analogue scale

The visual analogue scale (VAS) was a velcro-covered A4 plastic board with a 30cm vertical line and 10 intervals from 1-10 marked, as well as velcro-backed laminated cards (Figure 7). Each card was labelled with a SanQoL attribute and descriptor from the qualitative research:

1. Disgust - no smells or flies from the toilet
2. Health - the toilet doesn't spread diseases
3. Shame - I don't feel embarrassed because of the toilet
4. Safety - I feel safe using the toilet
5. Privacy - Nobody can see me in the toilet

Figure 7: SanQoL attributes visual analogue scale



Participants were first asked to choose the card representing the attribute they thought most important for a good bathroom and a good life. They were asked not to focus on their current bathroom but consider their ideal bathroom. They were then asked to do the same for the least important attribute of the remaining cards. These were placed at the top (10) and bottom (1) of the board. The enumerator explained that being at position 1 does not mean that attribute is not important, but just that it is *less* important than the others. The respondent was then asked to stick the remaining cards to the board, at the places on the line that they consider reflected relative importance. They were allowed to place more than one attribute at the same position, and to move attributes if they changed their mind. The exercise was based on methods reported in Drummond et al. (2015).

E. Reporting against ISOQOL criteria

Table 19 – ISOQOL minimum criteria (Reeve et al., 2013)

Area	Specific criterion	Comments and page references for this study
1. Conceptual and measurement model	Defining and describing the concept	p.128
	Intended population for use	p.127
	How the concept is organised into a measurement model	Figure 1
2. Reliability	Internal consistency reliability	p.142
	Test–retest reliability	p.142
3a. Content validity	Methods used to solicit and confirm attributes	p.128
	Characteristics of participants included in the evaluation	Table 3
	Justification for the recall period	p.132
3b. Construct validity	Empirical findings that support predefined hypotheses on expected associations	p.143
3c. Responsiveness	Evidence of changes in scores consistent with predefined hypotheses regarding changes	n/a - not longitudinal data
4. Interpretability of scores	What low and high scores represent for the measured concept	p.134
5. Translation of the measure	Methods used to translate and evaluate the measure	p.132
6. Patient and investigator Burden	Not overly burdensome for patients or investigators	Five items, with question length ranging from 8-11 words

Appendix D: Supplementary materials for Chapter 7 (evaluation paper)

A. Additional information on setting and intervention

B. Ranking exercise

C. Additional results and underlying data

D. Additional regression output

E. Robustness checks

A. Additional information on setting and intervention

A1 – Photographs with typical examples of main toilet types

Below are photographs of typical toilets of each type. CSB and ST designs are fairly homogenous, with some variation in the type of squat plate or seat pan used. Pit latrines are far more diverse. Some nominally meet the WHO/UNICEF Joint Monitoring Programme's definition of an improved technology (e.g. photo 2 has a concrete slab). These therefore be categorised as "limited" sanitation (since they are shared) rather than "unimproved".

Pit latrines (control)

1. Pit latrine with tyre and wood for squatting	2. Pit latrine with concrete slab
	
3. Fabric door providing limited privacy	4. No door and adjacent greywater pit
	

Shared toilets and community sanitation blocks (intervention)

Exterior	
1. Shared toilet (ST)	2. Community sanitation block (CSB)
	
Interior (varied between CSB / ST depending on design)	
3. Squat pan	4. Seat pan
	

Map of respondent households within Maputo

Panel A shows the greater Maputo region which, including the adjoining city of Matola, has a population of 2.9 million (INE, 2019). Panel B shows the geolocations of households included in our survey (n=424). They are situated within a small area of about 10km^2 within the Nhlamankulu district. Since compounds were randomly sampled from the list of MapSan-enrolled compounds, this broadly represents the implementation area of the intervention overall.

Figure 1: Maps of Maputo

A. Greater Maputo region

B. Respondent households within Maputo City

Source: Batran et al. (2018)

Source: Google Earth

Additional information on intervention delivery

The roles of key stakeholders involved in intervention delivery are summarised below. More information is provided in Brown et al. (2015) and Mattson (2016).

Table 20: Stakeholders involved in intervention delivery

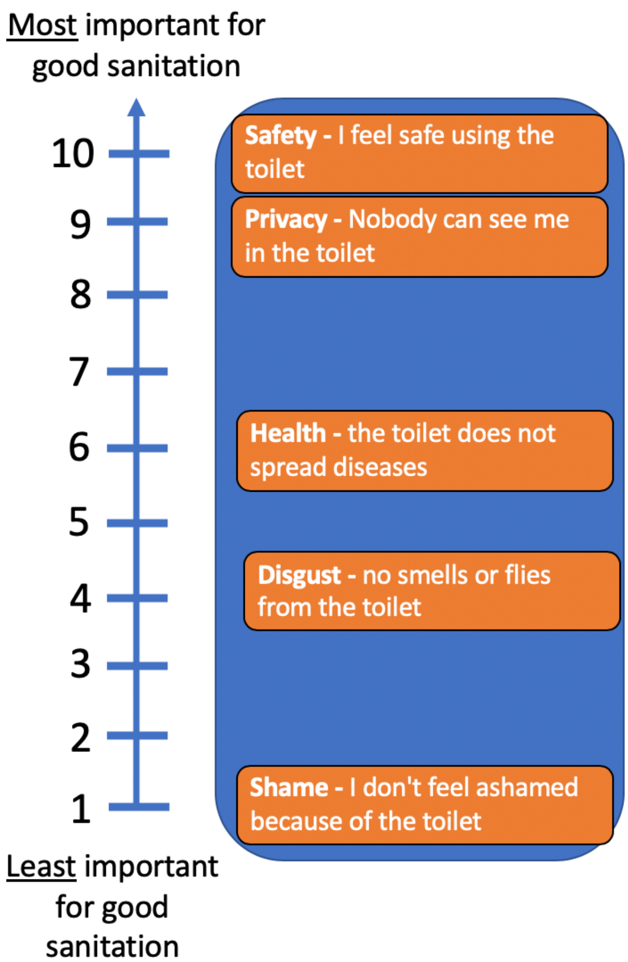
Stakeholder	Overall role	Key activities
WSUP (NGO)	Provider and project lead	• Project design and management • Manage design consultants • Manage construction contractors • Supervise construction
Community-based organisations	Sub-contractor facilitating community engagement	• Facilitate site selection • Collect household capital contribution
Construction firms	Sub-contractors constructing the toilet infrastructure	• Dismantle old toilet • Construct new toilet
Households	User of infrastructure	• Contribute 10-15% of capital costs • Clear site of material • Participate in meetings and data collection
Municipality (*Conselho Municipal de Maputo*, CMM)	Oversight and approvals by department for water and sanitation	• Approve designs and procurement • Provide permits for CSBs • Monitor infrastructure
World Bank	Oversight of overall programme	• Fund overall project • Oversight of delivery

B. Ranking exercise

The visual analogue scale (VAS) was a velcro-covered A4 plastic board with a 30cm vertical line and 10 intervals from 1-10 marked, as well as velcro-backed laminated cards (Figure 7). Each card was labelled with a SanQoL attribute and descriptor from the qualitative research:

1. Disgust - no smells or flies from the toilet
2. Health - the toilet doesn't spread diseases
3. Shame - I don't feel embarrassed because of the toilet
4. Safety - I feel safe using the toilet
5. Privacy - Nobody can see me in the toilet

Figure 2: SanQoL attributes visual analogue scale



Participants were first asked to choose the card representing the attribute they thought most important for a good bathroom and a good life. They were asked not to focus on their current bathroom but consider their ideal bathroom. They were then asked to do the same for the least important attribute of the remaining cards. These were placed at the top (10) and bottom (1) of the board. The enumerator explained that being at position 1 does not mean that attribute is not important, but just that it is *less* important than the others. The respondent was then asked to stick the remaining cards to the board, at the places on the line that they consider reflected relative importance. They were allowed to place more than one attribute at the same position, and to move attributes if they changed their mind. The exercise was based on methods reported in Drummond et al. (2015).

C. Additional results and underlying data

Figure 3: Histograms of primary and secondary outcomes by toilet type

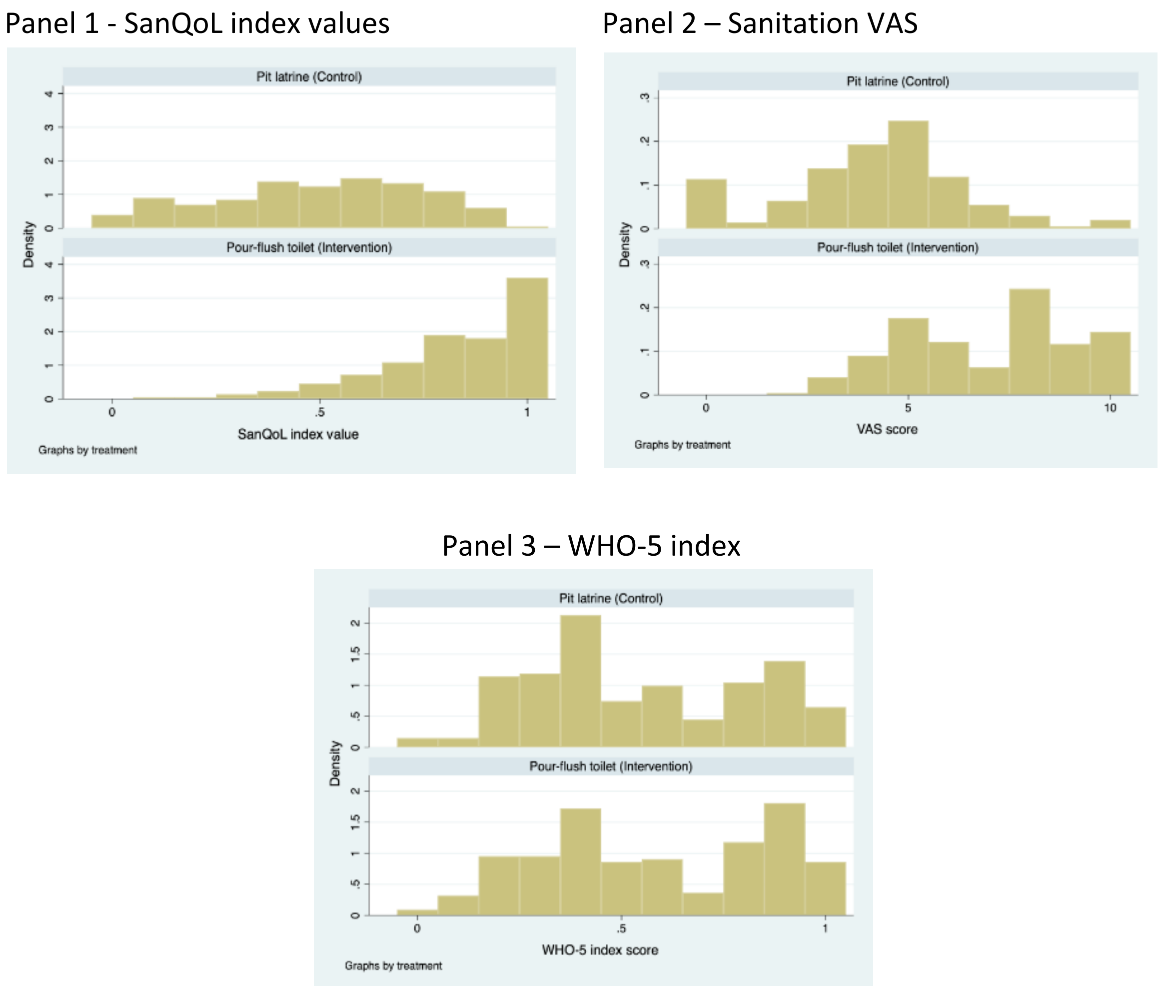
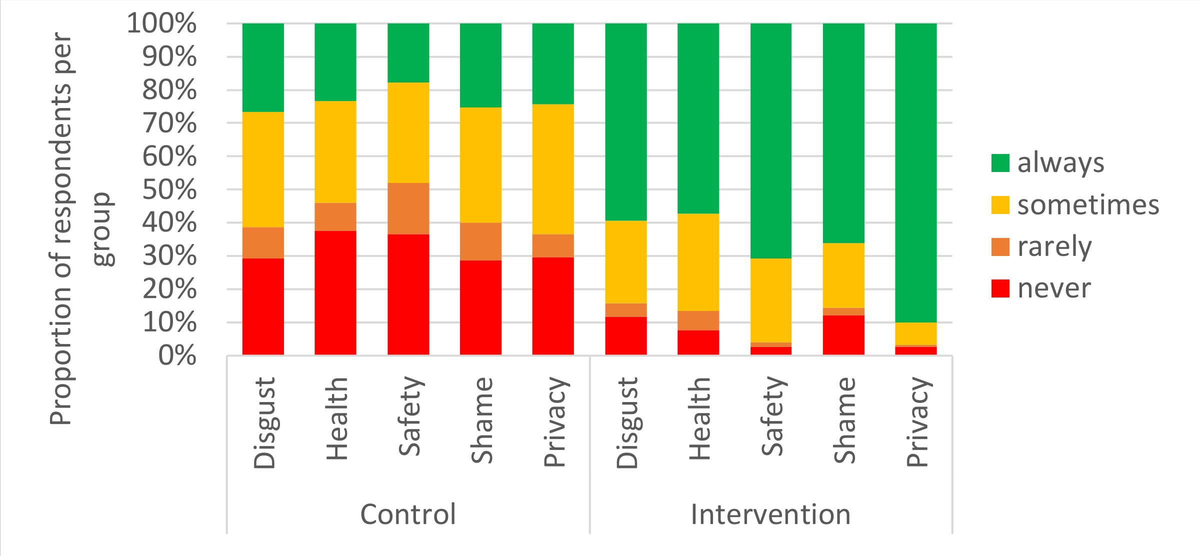


Figure 4: Distributions of SanQoL attributes by intervention and control



Note: Scores range from 0-3 representing a scale from never to always.

Table 21: Polychoric inter-item correlations for SanQoL attributes

	Disgust	Health	Shame	Safety	Privacy
Disgust	1.00				
Health	0.56	1.00			
Shame	0.52	0.53	1.00		
Safety	0.41	0.47	0.49	1.00	
Privacy	0.40	0.43	0.54	0.70	1.00

SanQoL questions in Portuguese

#	Dimension	Question	Responses
1	Disgust *Nojo*	Can you use the toilet without feeling disgusted? *Pode usar a casa de banho sem se sentir nojo?*	3 – Always (*sempre*) 2 – Sometimes (*as vezes*) 1 – Rarely (*raramente*) 0 – Never (*nunca*)
2	Health *Saúde*	Can you use the toilet without worrying that it spreads diseases? *Pode usar a casa de banho sem se preocupar que espalhe doenças?*	
3	Privacy *Privacidade*	Can you use the toilet in private, without being seen? *Pode usar a casa de banho com privacidade, sem ser visto/a?*	
4	Shame *Vergonha*	Can you use the toilet without feeling ashamed for any reason? *Pode usar a casa de banho sem sentir vergonha por qualquer motivo?*	
5	Safety *Segurança*	Are you able to feel safe while using the toilet? *É capaz de se sentir seguro/a ao usar esta casa de banho?*	

D. Additional regression output

Below is the regression output underlying adjusted models in Tables 4 and 5. Residuals for columns 1, 4, and 7 are plotted further below.

Table 22: Regression output underlying Tables 3 and 4

	SanQoL index value			**Sanitation-VAS**			**WHO-5 index**		
	(1) main regression	(2) gender interaction	(3) age interaction	(4) main regression	(5) gender interaction	(6) age interaction	(7) main regression	(8) gender interaction	(9) age interaction
Intervention toilet	0.339*** (0.023)	0.325*** (0.032)	0.334*** (0.024)	2.911*** (0.242)	2.758*** (0.305)	2.913*** (0.258)	6.20** (3.05)	6.60* (3.81)	6.74** (3.22)
Aged 60+	-0.012 (0.030)	-0.010 (0.031)	-0.031 (0.054)	-0.151 (0.270)	-0.120 (0.277)	-0.142 (0.453)	-12.93*** (2.90)	-13.01*** (2.95)	-10.60** (4.20)
Female	-0.007 (0.018)	-0.021 (0.031)	-0.006 (0.019)	-0.299* (0.157)	-0.452* (0.241)	-0.300* (0.160)	-3.31* (1.94)	-2.92 (2.56)	-3.47* (1.98)
Wealth index score	-0.004 (0.011)	-0.004 (0.011)	-0.004 (0.011)	-0.072 (0.100)	-0.072 (0.099)	-0.072 (0.100)	1.08 (1.12)	1.08 (1.12)	1.10 (1.12)
Intervention toilet * female		0.027 (0.039)			0.291 (0.327)			-0.75 (3.88)	
Intervention toilet * aged 60+			0.032 (0.064)			-0.016 (0.565)			-4.11 (5.68)
Constant	0.503*** (0.022)	0.510*** (0.026)	0.504*** (0.022)	4.284*** (0.200)	4.357*** (0.219)	4.669*** (0.407)	57.02*** (2.40)	56.84*** (2.53)	56.84*** (2.40)
Observations	423	423	423	423	423	423	422	422	422

Note: Cells report regression coefficients, with standard errors (clustered at compound level) in parentheses. *, **, *** indicate significance at the 10, 5 and 1 percent level. SanQoL is on a 0-1 scale, VAS is on a 0-10 scale, and WHO-5 is on a 0-100 scale.

1. meglm sqwt i.treatment i.over60 i.female wealth || c2a:, base vce(cluster c2a)

   where:
   - sqwt is SanQoL index values, a continuous variable ranging from 0-1
   - treatment is a dummy with value 1 for the intervention group
   - over60 is a dummy with value 1 if the respondent is aged 60 or over
   - female is a dummy variable with value 1 if the respondent is female
   - wealth is the wealth index score, coded as a continuous variable

2. meglm v1 i.treatment i.over60 i.female wealth || c2a:, base vce(cluster c2a)

   where:
   - v1 is sanitation VAS score, analysed as a continuous variable ranging from 0-10

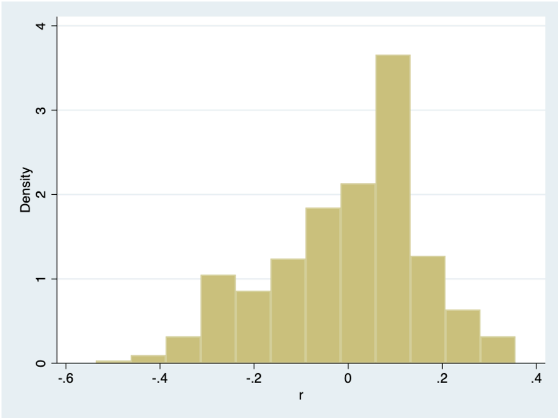
3. meglm who5index i.treatment i.over60 i.female wealth || c2a:, base vce(cluster c2a)

   where:
   - who5index is WHO-5 mental wellbeing, a continuous variable ranging from 0-1

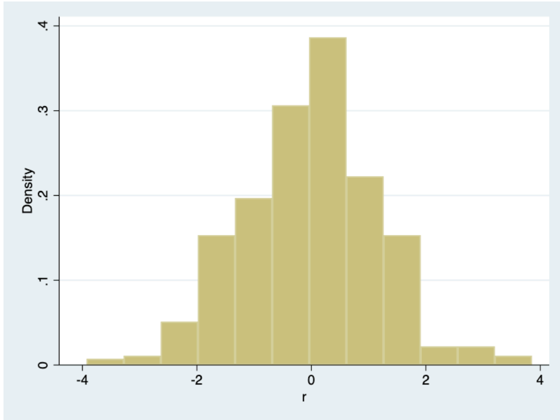
Diagnostic plots of residuals for the three main results (columns 1,4,7) are below. Histograms of residuals (Panels 1-3) show that they are approximately normally distributed for all three main results. Plots of residuals against fitted values for the fixed portion of the mixed model (Panels 4-6) raise no concerns about heteroscedasticity. For the SanQoL plot (panel 4) the plot for the intervention group appears truncated relative to control. This is a result of the modal SanQoL index value being 1 (see histogram in Supplementary Material C), which effectively censors the residuals.

Figure 5: Diagnostic plots of residuals for main results

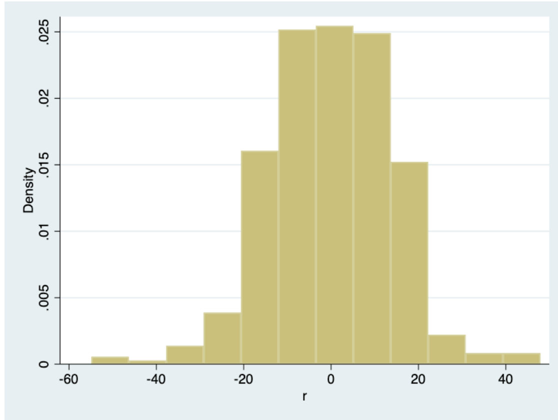
Panel 1 – Residuals for adjusted SanQoL index values regression (column 1)



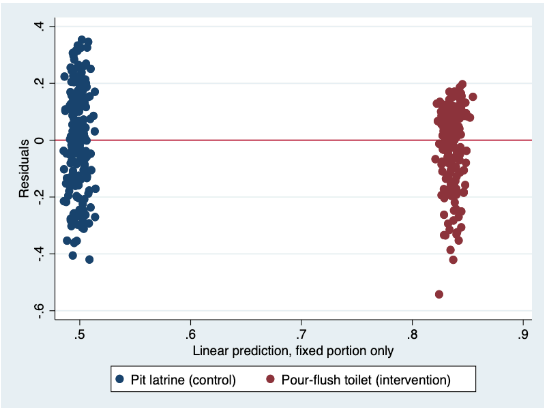
Panel 2 – Residuals for adjusted VAS regression (column 4)



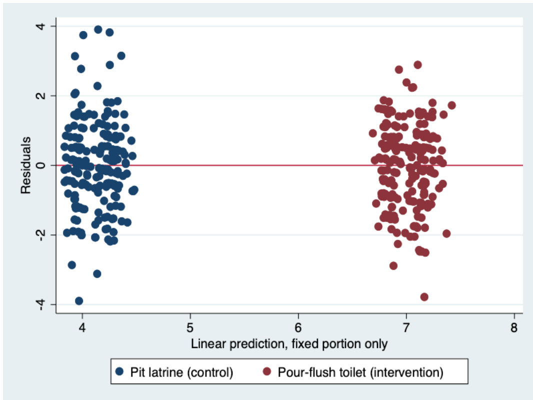
Panel 3 – Residuals for adjusted WHO-5 regression (column 7)



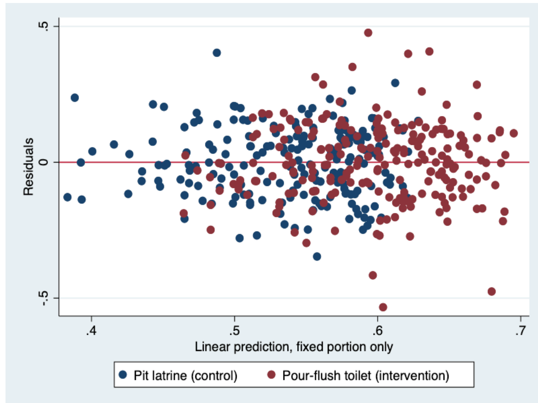
Panel 4 – Residuals against fitted for fixed portion – SanQoL (column 1)



Panel 5 – Residuals against fitted for fixed portion – VAS (column 4)



Panel 6 – Residuals against fitted for fixed portion – WHO-5 (column 7)



Below is the regression output for the main results regressions on SanQoL index values, but for users of shared toilets (ST) and community sanitation blocks (CSB) separately. Models are equivalent to column 1 in Table 4. Each is compared to the full control group. We include this for the purposes of subsequent cost-effectiveness analysis comparing these two options.

Table 23: Regression output for ST and CSB separately

	SanQoL index value	
	(1) Shared toilet (ST)	(2) Community sanitation block (CSB)
Intervention toilet	0.350*** (0.024)	0.282*** (0.040)
Aged 60+	-0.021 (0.031)	-0.016 (0.053)
Female	-0.007 (0.020)	-0.020 (0.028)
Wealth index score	-0.005 (0.011)	-0.000 (0.016)
Constant	0.504*** (0.022)	0.510*** (0.024)
Observations	385	240

Note: Cells report regression coefficients, with standard errors (clustered at compound level) in parentheses. *, **, *** indicate significance at the 10, 5 and 1 percent level

Below is the regression output for the regressions on individual SanQoL attributes, rather than combined index values. Further below is the results of similar regressions but including interaction terms. Together these results underlie Table 5.

Table 24: Regression output underlying Table 5 main results

	SanQoL attributes				
	Disgust	Health	Shame	Safety	Privacy
Intervention toilet	0.748*** (0.114)	0.963*** (0.112)	0.797*** (0.113)	1.360*** (0.100)	1.249*** (0.097)
Aged 60+	-0.069 (0.166)	-0.071 (0.134)	0.030 (0.154)	0.059 (0.127)	-0.096 (0.125)
Female	-0.037 (0.095)	0.080 (0.094)	0.074 (0.098)	-0.295*** (0.078)	0.021 (0.077)
Wealth index score	-0.104** (0.050)	-0.035 (0.052)	0.075 (0.052)	-0.001 (0.044)	0.034 (0.042)
Constant	1.624*** (0.100)	1.378*** (0.102)	1.550*** (0.098)	1.431*** (0.098)	1.591*** (0.097)
Observations	423	423	423	423	423

Note: Cells report regression coefficients, with standard errors (clustered at compound level) in parentheses. *, **, *** indicate significance at the 10, 5 and 1 percent level. Attribute scores are on a 0-3 scale.

Table 25: Regression output underlying Table 5 interactions

	SanQoL attributes: gender interactions					SanQoL attributes: age interactions				
	Disgust	Health	Shame	Safety	Privacy	Disgust	Health	Shame	Safety	Privacy
Intervention toilet	0.689*** (0.154)	0.961*** (0.153)	0.661*** (0.162)	1.267*** (0.129)	1.267*** (0.125)	0.715*** (0.122)	0.921*** (0.117)	0.781*** (0.119)	1.424*** (0.108)	1.239*** (0.104)
Aged 60+	-0.060 (0.170)	-0.071 (0.137)	0.052 (0.155)	0.074 (0.129)	-0.099 (0.127)	-0.220 (0.277)	-0.261 (0.240)	-0.045 (0.280)	0.343 (0.236)	-0.137 (0.224)
Female	-0.097 (0.153)	0.078 (0.158)	-0.064 (0.157)	-0.388*** (0.138)	0.039 (0.141)	-0.028 (0.097)	0.091 (0.096)	0.078 (0.099)	-0.314*** (0.078)	0.024 (0.078)
Wealth index score	-0.104** (0.050)	-0.035 (0.052)	0.074 (0.052)	-0.002 (0.044)	0.034 (0.042)	-0.105** (0.050)	-0.036 (0.052)	0.075 (0.052)	-0.001 (0.044)	0.034 (0.042)
Intervention toilet # female	0.114 (0.196)	0.005 (0.194)	0.262 (0.201)	0.177 (0.167)	-0.035 (0.159)					
Intervention toilet # aged 60+						0.262 (0.343)	0.331 (0.286)	0.131 (0.336)	-0.490* (0.270)	0.073 (0.265)
Constant	1.653*** (0.115)	1.379*** (0.118)	1.616*** (0.115)	1.475*** (0.112)	1.582*** (0.111)	1.636*** (0.101)	1.393*** (0.102)	1.556*** (0.099)	1.408*** (0.100)	1.594*** (0.099)
Observations	423	423	423	423	423	423	423	423	423	423

Note: Cells report regression coefficients, with standard errors (clustered at compound level) in parentheses. *, **, *** indicate significance at the 10, 5 and 1 percent level. Attribute scores are on a 0-3 scale.

E. Robustness checks

Table 26: Robustness checks for SanQoL and VAS

	Outcome: SanQoL index value (0-1 scale)					**Outcome: Sanitation VAS (0-10 scale)**				
	(1) Main model MEGLM (same as Table 4)	(2) Main model GEE	(3) Main model OLS	(4) MEGLM with only 10% level covariates	(5) MEGLM with all theory-based covariates	(6) Main model MEGLM (same as Table 3)	(7) Main model GEE	(8) Main model OLS	(9) MEGLM with only 10% level covariates	(10) MEGLM with all theory-based covariates
Pour-flush toilet (Intervention)	0.339*** (0.023)	0.339*** (0.023)	0.340*** (0.023)	0.337*** (0.023)	0.339*** (0.022)	2.911*** (0.242)	2.911*** (0.241)	2.867*** (0.246)	2.867*** (0.242)	2.895*** (0.241)
Aged 60+	-0.013 (0.030)	-0.013 (0.030)	-0.021 (0.031)		-0.010 (0.027)	-0.151 (0.270)	-0.153 (0.275)	-0.362 (0.279)		-0.143 (0.273)
Female	-0.007 (0.018)	-0.007 (0.019)	-0.005 (0.019)		-0.008 (0.018)	-0.300* (0.157)	-0.300* (0.160)	-0.261 (0.168)		-0.296* (0.158)
Wealth index score	-0.006 (0.011)	-0.006 (0.011)	-0.001 (0.010)	-0.007 (0.011)		-0.074 (0.100)	-0.074 (0.099)	-0.090 (0.112)	-0.092 (0.106)	
Participant age (continuous)				-0.000 (0.001)					-0.001 (0.006)	
Completed secondary school or above				0.046 (0.030)					0.608** (0.247)	
Number of people sharing toilet stance				-0.002 (0.002)	0.002 (0.003)				-0.004 (0.020)	0.024 (0.022)
Shares toilet with other households					-0.125*** (0.032)					-0.841*** (0.302)
Renter					0.027 (0.022)					0.023 (0.209)
Constant	0.503*** (0.022)	0.503*** (0.019)	0.498*** (0.022)	0.515*** (0.038)	0.572*** (0.035)	4.285*** (0.200)	4.285*** (0.192)	4.303*** (0.204)	4.162*** (0.386)	4.687*** (0.352)
Observations	423	423	423	423	424	423	423	423	423	424

Note: standard errors are shown in parentheses, which are clustered at the compound level; *, **, *** indicate significance at the 10, 5 and 1 percent level

Table 27: Robustness checks for WHO-5

	(1) Main model MEGLM (same as Table 4)	(2) Main model GEE	(3) Main model OLS	(4) MEGLM with only 10% level covariates	(5) MEGLM with all theory-based covariates	(6) MEGLM with health determinant covariates
Pour-flush toilet (Intervention)	6.246** (3.051)	6.272** (3.044)	4.833 (3.153)	6.384** (3.018)	5.985** (3.035)	6.214** (2.882)
Aged 60+	-12.951*** (2.906)	-12.867*** (3.298)	-15.975*** (3.414)		-13.038*** (2.901)	-6.195** (3.131)
Female	-3.314* (1.940)	-3.344* (1.878)	-2.187 (2.057)		-3.664* (1.931)	-1.295 (2.010)
Wealth index score	0.982 (1.118)	0.988 (1.191)	0.785 (1.358)	0.650 (1.164)		
Participant age (continuous)				-0.324*** (0.072)		
Completed secondary school or above				1.420 (3.589)		
Number of people sharing toilet stance				0.248 (0.268)	0.460* (0.276)	
Shares toilet with other households					-5.456 (4.129)	
Renter					-1.329 (2.758)	
Has partner						-1.954 (2.094)
Pain scale						-7.233*** (2.239)
Problems walking scale						-5.789** (2.706)
Constant	57.007*** (2.398)	57.001*** (2.394)	57.391*** (2.470)	63.475*** (4.802)	56.672*** (4.479)	73.259*** (4.061)
Observations	422	422	422	422	423	423

Standard errors are shown in parentheses, which are clustered at the compound level; *, **, *** indicate significance at the 10, 5 and 1 percent level

Table 28: Robustness checks for attribute level models (coefficients are odds ratios)

	Meologit for main attribute-level model				
	Disgust	Health	Shame	Safety	Privacy
Intervention toilet	5.58***	7.94***	5.51***	26.75***	65.66***
	(1.70)	(2.52)	(1.46)	(11.15)	(41.82)
Aged 60+	0.93	0.78	1.08	1.08	0.83
	(0.34)	(0.24)	(0.36)	(0.39)	(0.30)
female	0.94	1.22	1.09	0.45***	1.07
	(0.19)	(0.25)	(0.22)	(0.10)	(0.26)
wealth index score	0.80*	0.91	1.16	0.99	1.08
	(0.10)	(0.11)	(0.12)	(0.11)	(0.14)
Observations	423	423	423	423	423

Note: coefficients are odds ratios. Standard errors are shown in parentheses, which are clustered at the compound level; *, **, *** indicate significance at the 10, 5 and 1 percent level

Table 29: Robustness checks for attribute level interactions (coefficients are odds ratios)

	Meologit for attribute-level gender interaction					Meologit for attribute-level age interaction				
	Disgust	Health	Shame	Safety	Privacy	Disgust	Health	Shame	Safety	Privacy
Intervention toilet	4.87***	7.58***	4.42***	29.95***	71.50***	5.28***	7.36***	5.33***	35.60***	57.12***
	(1.83)	(2.97)	(1.52)	(16.75)	(50.93)	(1.69)	(2.42)	(1.44)	(16.98)	(36.37)
Aged 60+	0.95	0.79	1.11	1.06	0.82	0.75	0.57	0.95	2.11	0.68
	(0.36)	(0.25)	(0.37)	(0.39)	(0.31)	(0.41)	(0.28)	(0.47)	(1.02)	(0.28)
Female	0.83	1.17	0.90	0.48***	1.12	0.96	1.24	1.10	0.41***	1.09
	(0.24)	(0.36)	(0.23)	(0.13)	(0.31)	(0.20)	(0.26)	(0.22)	(0.10)	(0.26)
Wealth index score	0.80*	0.91	1.16	0.99	1.08	0.80*	0.91	1.16	0.98	1.09
	(0.10)	(0.11)	(0.12)	(0.11)	(0.14)	(0.10)	(0.11)	(0.12)	(0.11)	(0.14)
Female*Intervention interaction	1.29	1.09	1.54	0.84	0.86					
	(0.54)	(0.46)	(0.63)	(0.39)	(0.49)					
Age*Intervention interaction						1.49	1.74	1.29	0.26*	2.06
						(1.12)	(1.11)	(0.89)	(0.18)	(2.22)
Observations	423	423	423	423	423	423	423	423	423	423

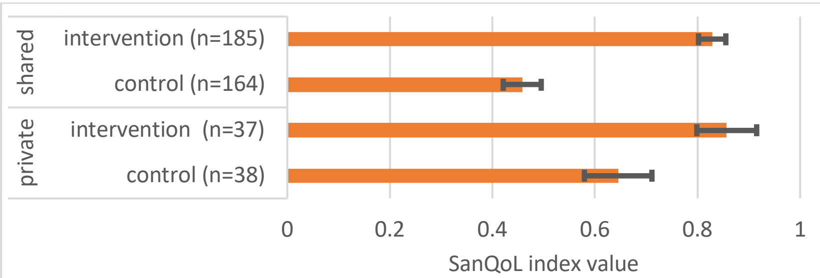
Note: coefficients are odds ratios. Standard errors are shown in parentheses, which are clustered at the compound level; *, **, *** indicate significance at the 10, 5 and 1 percent level

The role of sharing toilets

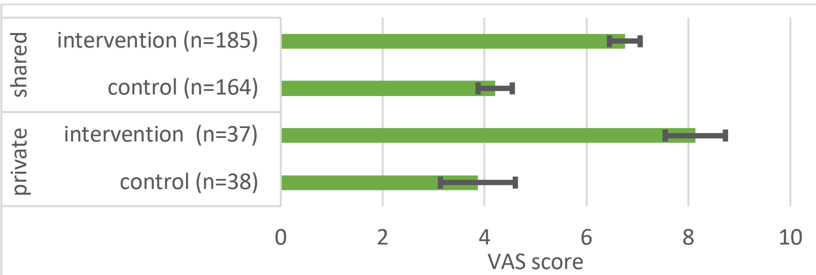
During study design it was anticipated that most people in our sample would be sharing toilets with other households, as they had been at MapSan baseline four years previously. However, this was not a sampling criterion. In the event, 81% of control and 83% of intervention households used shared toilets. The households using private toilets were single-household compounds, likely due to empty dwellings (driven by rental markets or migration) or changes in compound living arrangements in the four years since the intervention. It was more surprising that 37 out of 185 intervention compounds (20%) sampled now had toilets being used by only one household. There was evidence for mean values of our three outcomes per treatment group varying according to sharing status (Figure 6).

Figure 6: Differences between groups using private and shared toilets

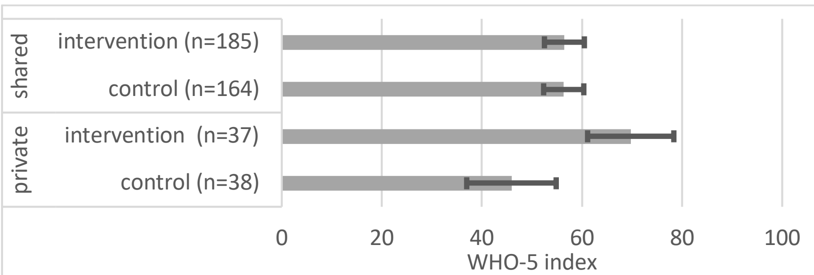
(a) Mean SanQoL index value by sharing and treatment group



(b) Mean VAS score by sharing and treatment group



(c) Mean WHO-5 index by sharing and treatment group



After discussing mean values in these data, we explore them in a further set of regressions. In the set of robustness checks including all covariates hypothesised *ex ante* as influencing SanQoL, the binary covariate for sharing the toilet with other households was significant at the 1% level with a negative coefficient. This is explained by differences within the control group, where people using private PLs had substantially higher SanQoL than people using shared PLs (Figure 6a). In the intervention group, by contrast, sharing made little difference to SanQoL. Considering VAS scores (Figure 6a), a slightly different pattern was observed, whereby mean scores amongst people sharing intervention toilets were lower than those not sharing. By contrast, sharing made no difference within the control group. For both outcomes however, the intervention was associated with a substantial difference regardless of sharing status. For WHO-5, the picture is different again. The intervention made a substantial difference to WHO-5 amongst people using private toilets, while there was no difference amongst people sharing. Note that The mean number of people sharing stances of shared or private toilets did not significantly differ between intervention and control groups (5.9 and 6.3 for private toilets in control and intervention groups respectively, compared to 12.4 and 13.3 for shared toilets).

We ran a regression specified according to the headline results but including a factorial interaction term between the intervention and the binary sharing covariate (Table 30). The results are easier to interpret in the light of Figure 6. We make three interpretations from these results. First, amongst people sharing toilets with other households, there was very strong evidence ($p<0.001$) that the intervention was associated with a difference in SanQoL of 0.38 (95% CI: 0.33 – 0.43). This is greater than the difference of 0.34 in the sample as a whole, and substantially larger than the difference of 0.19 (95% CI: 0.10 – 0.28) amongst those using private toilets. Second, the opposite trend was seen in VAS scores: amongst users of shared toilets the intervention was associated with a difference of 2.7 (95% CI: 2.1 – 3.2), while in private toilets it was associated with a difference of 4.0 (95% CI: 3.0 – 4.9). Third, for WHO-5 scores, there was no evidence of a difference amongst users of shared toilets (95% CI: -4.0 – 9.3), compared to a substantial difference of 19.9 amongst users of private toilets (95% CI: 6.6 – 33.2).

The mixed results for the three outcomes frustrate attempts at a combined interpretation.

SanQoL was the primary outcome of our study, and we found a bigger difference amongst people sharing toilets than people using private toilets. At time of delivery, the intervention was targeted at people using shared toilets, and aimed to deliver high-quality shared sanitation rather than private sanitation. This invites the conclusion that the main results of interest should be the overall findings as well as the findings in this section for the shared sub-group. Less attention should be paid to comparisons with the private sub-group, since it was small. However, it is important that only four years after the intervention, the benefits of toilets which were meant to be shared were in fact being enjoyed by only one household (with mean size 6) in 20% of intervention compounds. Our study provides good evidence that access to shared pour-flush toilets by comparison to shared pit latrines were associated with a substantial improvement in SanQoL index values and VAS scores, but no difference in mental wellbeing. To explore the quality of life effects of shared sanitation by comparison to private toilets, future studies would need to be adequately powered.

Table 30: Interactions with by sex and age-group

	SanQoL index value	VAS score	WHO-5 index score
Intervention	0.19*** (0.05)	3.95*** (0.49)	19.88*** (6.78)
Aged 60+	-0.02 (0.03)	-0.15 (0.27)	-12.81*** (2.91)
Female	-0.01 (0.02)	-0.32** (0.16)	-3.48* (1.96)
Wealth index score	-0.01 (0.01)	-0.09 (0.10)	0.85 (1.11)
Shares toilet with other households	-0.20*** (0.04)	-0.01 (0.45)	5.68 (5.31)
Intervention # shares toilet	0.19*** (0.05)	-1.30** (0.56)	-17.23** (7.63)
Constant	0.67*** (0.04)	4.30*** (0.41)	52.61*** (5.00)
Observations	423	423	422

Appendix E: Supplementary materials for Chapter 9 (cost-effectiveness paper)

A. Additional intervention delivery information and photos

B. Additional methodological information

C. Additional costing methods

D. Additional results

E. Reporting against CHEERS

A. Additional intervention delivery information and photos

Activities undertaken by each stakeholder are summarised below, covering the construction phase.

Table 1: Stakeholders involved in intervention delivery

Stakeholder	Overall role	Key activities
WSUP	Provider and project lead	• Project design and management • Manage contractors • Supervise construction
CBOs	Sub-contractor facilitating community engagement	• Facilitate site selection • Collect household capital contribution
Construction firms	Sub-contractors constructing the toilet infrastructure	• Dismantle old toilet • Construct new toilet
CMM	CMM department for water and sanitation	• Approve designs and procurement • Get permits for CSBs • Monitor infrastructure
World Bank	Oversight of overall programme	• Contribute c.10% of capital costs • Clear site of material • Participate in meetings and data collection
Households	User of infrastructure	• Contribute 10-15% of capital costs • Clear site of material • Participate in meetings and data collection

Pit latrines

1. Pit latrine with tyre and wood for squatting	2. Pit latrine with concrete slab
	

3. Fabric door providing limited privacy	4. No door and adjacent greywater pit
	

Shared toilets and community sanitation blocks

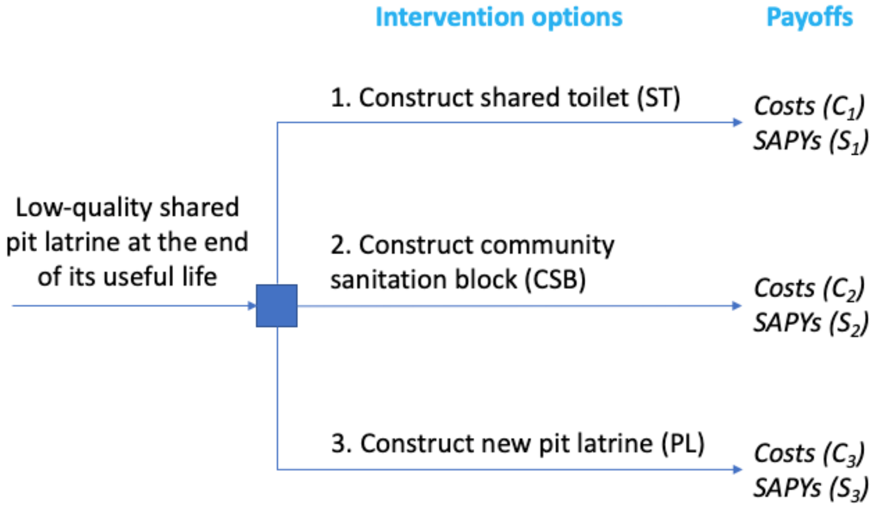
<table>
<tr><th colspan="2">Exterior</th></tr>
<tr><td>1. Shared toilet (ST)</td><td>2. Community sanitation block (CSB)</td></tr>
<tr><td></td><td></td></tr>
<tr><th colspan="2">Interior (varied between CSB / ST depending on design)</th></tr>
<tr><td>3. Squat pan</td><td>4. Seat pan</td></tr>
<tr><td></td><td></td></tr>
</table>

B. Additional methodological information

Model structure

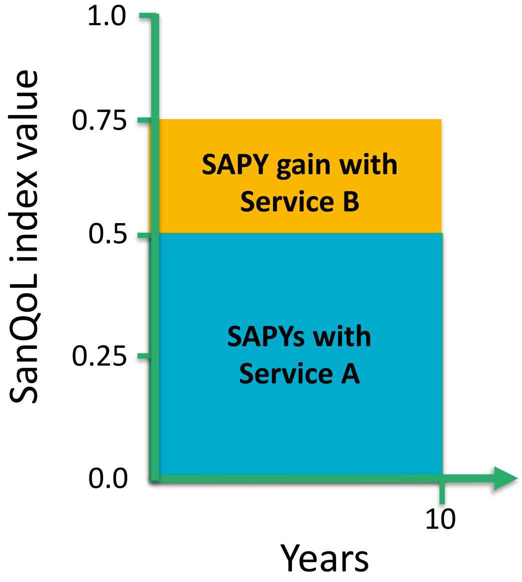
The decision tree comprising our decision analytic model is in Figure 7. Note that the options are mutually exclusive and no probabilities are modelled, so the pathway probability for each option is 1 once the decision has been made.

Figure 7: Decision tree model structure



Visualised example of a SAPY

A toilet used by five people all at full sanitation capability would generate five SAPYs per year. The figure provides a visual explanation in which in which two hypothetical sanitation technologies are compared. Both have a useful life of 10 years, but Service A provides the user a SanQoL of 0.5 while Service B provides 0.75. Service A generates 0.5 x 10 = 5 SAPYs to each user while Service B generates 0.75 x 10 = 7.5 SAPYs, representing a 'SAPY gain' of 2.5 over a 10-year horizon. Discounting is ignored for ease of explanation. A decision-maker would weigh up the incremental effect against the incremental cost of Service A, specifically the cost per SAPY gained.



Incremental cost-effectiveness ratios

We made three pair-wise comparisons. The incremental cost-effectiveness ratio (ICER) was calculated as follows for the ST option:

$$ICER\ (per\ SAPY\ gained) = \frac{C_{ST} - C_{pit}}{SAPY_{ST} - SAPY_{pit}}$$

C_{ST} and C_{pit} are total discounted economic costs in the respective groups, over the 15-year time horizon. $SAPY_{ST}$ and $SAPY_{pit}$ are the number of discounted SAPYs generated for the shared toilet and pit latrine options, respectively. We also calculated ICERs for the CSB as compared to the PL, and for the ST as compared to the CSB.

C. Additional costing methods

Costs estimated top-down

Table 2 presents estimation types and data sources for each of the cost categories in Table 3. Here we provide more information on methods for estimating some of these costs. For the ST and CSB options, categories representing about two thirds of lifecycle economic costs were estimated top-down from the NGO expenditure report to the World Bank (#1, #4, #5 and #9). Toilet construction was only one component of the broader project, which included activities supporting faecal sludge management service providers, for example.

First, based on interviews with NGO staff, we identified expenditure categories which contributed to toilet construction outputs. These could be direct, such as toilet construction contracts, or indirect, such as a share of overall monitoring and evaluation (M&E) costs. Allocation of some categories could be estimated with reasonable certainty based on interviews (e.g. NGO staff time). Other categories (e.g. M&E) required an allocation "by value", that is, based on the expenditure share of the toilet component by comparison to the other components. In this way, we estimated overall costs attributable to the toilet construction component.

Second, we allocated those attributable costs between STs and CSBs, again relying on interviews with NGO staff wherever possible. Category descriptions were not always clear, and interviews often allowed 100% of a line item to be allocated to STs, for example. We again used allocation by value when there was no reasonable basis for estimation. In this way, we estimated the NGO-borne costs of serving people with 400 STs and 50 CSBs. Based on the mean number of users per toilet type from the 2018 household survey, we aggregated those costs to the hypothetical cohort of 7,200.

Table 2: Costing methods and data sources per cost type

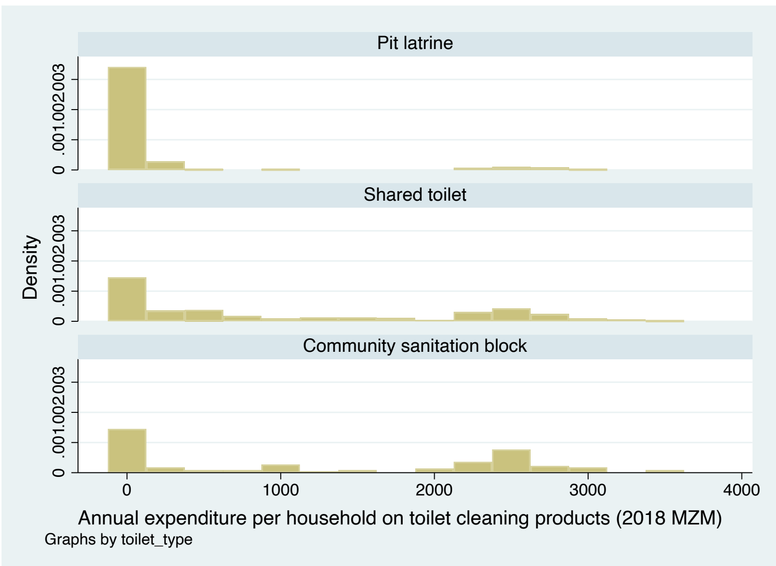
	Cost type	Estimation type		Data sources	
		ST and CSB	**PL**	**Quantities**	**Prices**
Capital	1. NGO staff	top-down	n/a	NGO expenditure report to WB	
	2. WB staff	bottom-up	n/a	WB timesheet records	WB salary data
	3. CMM staff	bottom-up	n/a	Interviews	CMM salary data
	4. NGO consultants	top-down	n/a	NGO expenditure report to WB	
	5. WB consultants	bottom-up	n/a	WB timesheet records	WB salary data
	6. Works	top-down	bottom-up	ST & CSB: NGO expenditure report to WB PL: quantities from household survey users per toilet, prices from household survey (expenditure last time constructed a PL)	
	7. Time of *Chefe do Quarteirão*	bottom-up	n/a	Interviews	opportunity cost of time (voluntary position)
	8. Household participation time	bottom-up	n/a	Interviews	opportunity cost of time
	9. NGO overheads		n/a	NGO expenditure report to WB	
Recurrent	10. Cleaning expenditure	bottom-up	bottom-up	household survey (% of households reporting incurring)	household survey (average expenditure amongst households incurring)
	11. Cleaning value of time	bottom-up	bottom-up		
	12. Maintenance expenditure	bottom-up	bottom-up		
	13. Emptying expenditure	bottom-up	bottom-up		
	14. Treatment cost	bottom-up	n/a	m3 of faecal sludge accumulated per person (Strande et al., 2018)	operational cost per m3 treated for faecal sludge treatment plants in Maputo Master Plan (AIAS, 2015)

Costs estimated bottom-up

For CMM staff (#3) and World Bank staff and consultants (#2,#5), we obtained individual-level salary data from these institutions. World Bank staff are paid net salaries with a tax allowance, so we assumed 30% uplift for tax allowance and pension (World Bank, 2019d). CMM salary data were provided disaggregated by base salary, bonuses, and fuel subsidies. We interviewed staff members involved in intervention delivery about the nature of their involvement. The World Bank had records of estimated time spent per activity, which we refined based on interviews. With CMM, we estimated time spent per activity based on interviews. The time of households (#8) for participating in the intervention (e.g. meetings), as well as for *Chefes do Quarteirão* (#7, informal volunteer neighbourhood leaders) was estimated via interviews with NGO staff. Their time was valued at 50% (Whittington and Cook, 2018) of the mean minimum wage in Mozambique in 2015. Minimum wages in Mozambique vary by industry - we calculated the mean minimum wage across all 15 industries for which data was available (WageIndicator, 2016).

Recurrent costs (#10-14) were all estimated bottom-up, mostly via the 2018 household survey described elsewhere (Capone et al., 2020). Taking the example of expenditure on cleaning, we asked respondents whether their household purchased any of six cleaning materials to be used primarily for cleaning the toilet: bleach, detergent, bucket, mop, broom, or any other purchases. For any they answered yes, we asked about the frequency with which it was purchased, and the approximate price last time. This allowed us to build a bottom-up picture of annual cleaning expenditure. Data for this variable are presented in Figure 8. As is common with cost data, the data are strongly right-skewed and there are many zero values – e.g. 56% of pit latrine users incurred no cleaning expenditure. This is why we used a gamma distribution for these variables in the PSA. We applied a similar approach to the value of cleaning time, maintenance expenditure and emptying expenditure.

Figure 8: Histograms of annual household expenditure on cleaning products, by toilet type



For emptying costs, since few households had emptied septic tanks in the 2-3 years since construction, we assumed that tanks would be emptied every 5 years on average. We therefore applied, to all individuals, mean costs reported by those households who had already incurred them. We assumed that prices paid for emptying services in private markets reflected the opportunity costs of both emptying and transport.

To estimate treatment costs, we first estimated a recurrent cost per m^3 for faecal sludge treatment plants in the 2015 Maputo sanitation master plan (AIAS, 2015). We took the mean operation and maintenance (O&M) cost per m^3 of sludge treated of the three proposed plants, for the medium-term scenario in the master plan. We then took the total faecal sludge accumulation rate for septic tanks reported by Strande et al. (2018) based on an empirical study in Uganda . It is reported in litres per person per year, so we converted this to a m^3 / year value for the 7,200 people in our cohort, and multiplied this by the master plan (O&M) cost.

D. Additional results

Affordability

We discuss the affordability of the ST intervention with respect to the budget of the *Conselho Municipal de Maputo* (CMM), the municipality (Bilinski et al., 2017). CMM recently approved a sanitation tax levied on water bills (Acey et al., 2019). Only sanitation interventions are eligible and the budget is fixed in the medium-term (CMM, 2017). CMM's annual infrastructure budget for water supply and sanitation was about $3 million in 2017 (CMM, 2018). Since water supply is the responsibility of another agency, the budget is mostly spent on drainage (55%) and sewerage (22%), with only 3% dedicated to non-sewered sanitation. If the non-sewered sanitation budget were increased to 10% of the total, that would amount to $300,000. Of Maputo's 1.2 million population, about 20% use pit latrines, i.e. 240,000 people (Hawkins and Muximpua, 2015). Many of these pit latrines will be private, rather than shared. Assuming that 100,000 people are sharing low-quality pit latrines, about 8,000 STs at the levels of sharing seen in our study would be required.

With an annual budget of $300,000, and maintaining subsidy at 85%, it would take 23 years to support all shared PL users to upgrade to shared ST. In short, the intervention is unaffordable within a realistic timeframe (Bilinski et al., 2017). This takes only works costs into consideration, which likely underestimates the true cost of a CMM-managed programme. However, costs may be lower than the total economic cost presented here, since NGO engineering designs could be reused and salaries for management roles would likely be lower. Reducing the subsidy to 50% would allow more users to be supported, but may not be affordable to households likely to be using low-quality PLs at present. The gap might be filled by international aid.

Disaggregated cost data

Table 3: Further disaggregated economic costs (in 2015 US dollars) of the three strategies

		Strategies					
		Pit latrine (PL)		Shared toilet (ST)		Community sanitation block (CSB)	
Cost categories		$	%	$	%	$	%
Capital	**NGO staff**						
	NGO staff in Mozambique			357,411	20%	404,535	10%
	NGO staff in UK			26,184	1%	29,636	1%
	WB staff						
	WB staff in Mozambique			38,649	2%	38,649	1%
	CMM staff						
	CMM staff			15,186	1%	15,186	0.4%
	NGO consultants						
	Consultants for site selection			72,833	4%		0.0%
	Consultants for design & supervision			70,515	4%	766,490	20%
	Consultants for monitoring & evaluation			67,016	4%	75,852	2%
	WB consultants						
	WB consultants (for monitoring)			11,220	1%	11,220	0.3%
	Works						
	Works (desludging old latrines)			20,146	1%	9,130	0.2%
	Works (toilet construction)	30,866	12%	441,134	24%	1,556,471	40%
	Time of Chefe do Quarteirão						
	Time of *Chefe do Quarteirão*			105	0.01%	105	0.003%
	Household participation time						
	Household participation time			3,913	0.2%	3,913	0.1%
	NGO overheads						
	training and workshops			16,982	1%	19,221	0.5%
	office/consumables			23,138	1%	26,189	1%
	transport			4,232	0.2%	4,790	0.1%
	UK running cost contribution			18,193	1%	20,592	1%
	audit contribution			20,199	1%	22,863	1%
	Capital sub-total	**30,866**	**12%**	**1,207,056**	**67%**	**3,004,841**	**77%**
Recurrent	Cleaning expenditure	94,335	37%	410,270	23%	618,768	16%
	Cleaning value of time	90,901	36%	117,320	6%	119,344	3%
	Maintenance expenditure	8,120	3%	13,602	1%	124,517	3%
	Emptying expenditure	29,802	12%	37,163	2%	27,744	1%
	Treatment cost			21,580	1%	21,580	1%
	Recurrent sub-total	**223,158**	**88%**	**599,935**	**33%**	**911,952**	**23%**
Total		**254,024**	**100%**	**1,806,992**	**100%**	**3,916,793**	**100%**

Further cost metrics

Table 4 presents unit cost metrics other than the headline results. The rows for works cost for example (#1), present how much it cost simply to build the toilet, excluding the associated programme costs. This type of cost metric is more often used in WASH cost analyses, but presents a partial picture of how much it costs to deliver the services. The capital cost (#2) shows how much the full cost of ensuring the infrastructure is available, without accounting for recurrent costs. These results may be useful for planning or for comparing to similar interventions. For both works and capital (#1 and #2) the ST costs about 40 times more per person than PL, and the CSB 100-130 times more than PL. These ratios show that the PL and CSB represent a substantial step up from PL in terms of capital cost. Users are unlikely to be willing to pay the full amount.

The recurrent cost (#3) shows both the total recurrent cost over the 15 year time horizon, as well as the average annual recurrent cost. The majority of these costs are borne by users. Recurrent costs per person are about three times more for STs than PLs, and about four times more for CSBs than PLs.

Table 4: Alternative unit cost metrics

		Pit latrine (PL)	**Shared toilet (ST)**	**Community sanitation block (CSB)**	**ST:PL ratio**	**CSB:PL ratio**
1. Cost of works (undiscounted year 0 financial cost)	works cost per toilet (construction contracts only)	20*	841	6,299	42	315
	works cost per person (construction contracts only)	1.6*	64	217	40	136
2. Capital cost (discounted economic cost)	capital cost per toilet (over 15 years)	53	2,201	12,090	42	228
	full capital cost per person (over 15 years)	4.3	168	417	39	97
3. Recurrent cost (discounted economic cost)	total recurrent cost per toilet (over 15 years)	383	1,094	3,669	3	10
	average annual recurrent cost per toilet	26	73	245	3	10
	total recurrent cost per person (over 15 years)	31	83	127	3	4

		Pit latrine (PL)	Shared toilet (ST)	Community sanitation block (CSB)		ST:PL ratio	CSB:PL ratio
	average annual recurrent cost per person	2	6	8		3	4
4. Total cost (discounted economic cost)	total cost per toilet (over 15 years)	436	3,295	15,759		8	36
	total cost per person (over 15 years) **	35	251	544		7	15

*for one toilet - analyses over 15 years in rest of column includes three toilet builds, since the useful life is 5 years. ** same as Table 3

Results in Mozambican meticais (MZN)

Table 5: Data for Table 3 presented in MZN

Lifecycle economic costs		Pit latrine (PL)		Shared toilet (ST)		Community sanitation block (CSB)	
		MZN	**%**	**MZN**	**%**	**MZN**	**%**
Capital	NGO staff			15,381,834	21%	17,409,894	11%
	WB staff			1,549,785	2%	1,549,785	1%
	CMM staff			608,944	1%	608,944	0.4%
	NGO consultants			8,435,418	12%	33,777,220	22%
	WB consultants			449,913	1%	449,913	0%
	Works	1,237,697	12%	18,496,946	26%	62,779,281	40%
	Time of *Chefe do Quarteirão*			4,228	0.006%	4,228	0.003%
	Household participation time			156,911	0.2%	156,911	0.1%
	NGO overheads			3,317,978	5%	3,755,446	2%
	Capital sub-total	**1,237,697**	**12%**	**48,401,955**	**67%**	**120,491,622**	**77%**
Recurrent	Cleaning expenditure	3,782,772	37%	16,451,487	23%	24,812,069	16%
	Cleaning value of time	3,645,059	36%	4,704,448	6%	4,785,575	3%
	Maintenance expenditure	325,587	3%	545,440	1%	4,993,022	3%
	Emptying expenditure	1,195,032	12%	1,490,188	2%	1,112,518	1%
	Treatment cost			865,339	1%	865,339	1%
	Recurrent sub-total	**8,948,450**	**88%**	**24,056,902**	**33%**	**36,568,522**	**23%**
Total		**10,186,147**	**100%**	**72,458,857**	**100%**	**157,060,144**	**100%**

Table 6: Data for Table 4 presented in MZN

		Pit latrine (PL)	Shared toilet (ST)	Community sanitation block (CSB)
Outputs	number of toilets	583	548	249
	mean users per toilet	12	13	29
	total people served	7,200	7,200	7,200
Outcomes*	total SAPYs	42,392	72,322	67,343
Total economic cost*	overall	10,186,147	72,458,857	157,060,144
	of which capital	1,237,697	48,401,955	120,491,622
	of which recurrent	8,948,450	24,056,902	36,568,522
Economic cost per person*	total cost per person	1,415	10,064	21,814
	annualised total cost per person	119	843	1,827
Cost-effectiveness	incremental cost effectiveness ratio (cost per SAPY gained, compared to PL)	.	2,081	dominated**

note. all costs in 2015 Mozambican meticais
* Summed over the 15-year time horizon and discounted
**Since CSB is dominated by ST, no ICER for CSB is on the cost-effectiveness frontier (Figure 1). In the specific circumstance where ST is infeasible on engineering grounds, for example, the ICER for CSB would be MZN 5,886 (incremental to PL)

Additional deterministic sensitivity analysis

We tested the same DSA scenarios as in Table 2 for the other two comparisons, excluding those which were not applicable (Table 7). The only result of note is that CSB was no longer dominated in one scenario in the ST-CSB comparison, yellow-highlighted in Table 7. This scenario is also presented in Figure 9.

Figure 9: Cost-effectiveness plane showing cost-effectiveness frontiers in the base case and the DSA scenario in which 95% CIs for SanQoL are at their upper bound for CSB and lower bound for ST

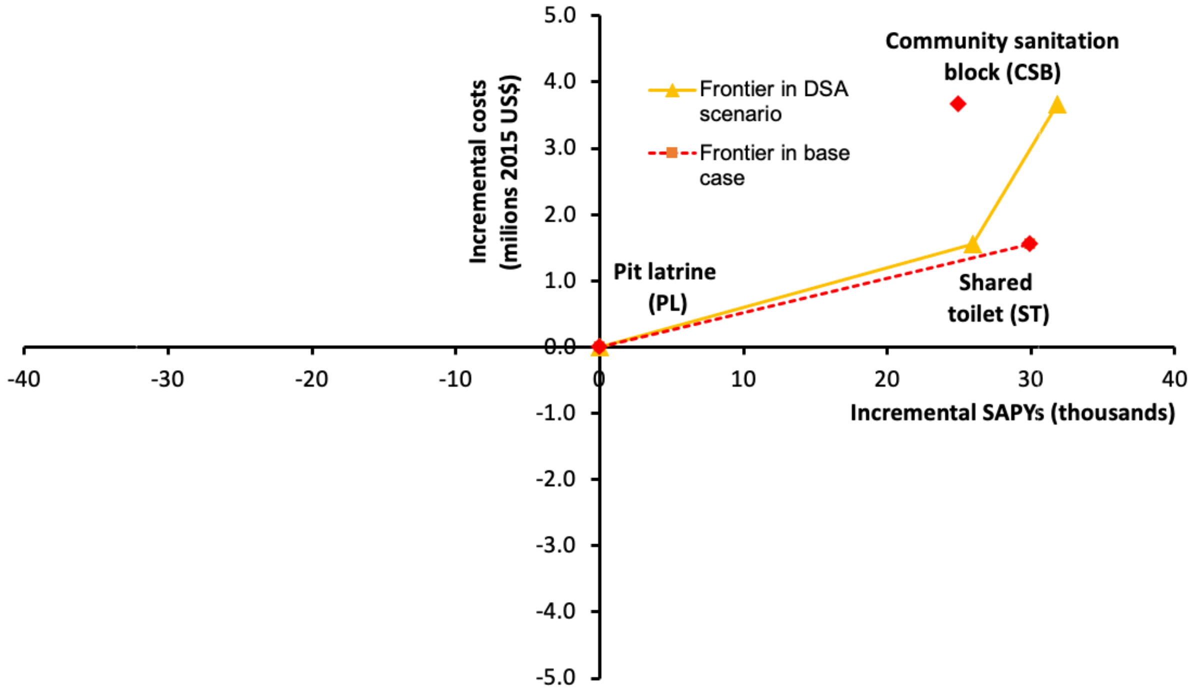


Table 7: Deterministic sensitivity analysis for other two scenarios

		Base case		-424		147	
	Parameter values			ST vs CSB		CSB vs PL	
Parameter	**Base case**	**Low**	**High**	**Low**	**High**	**Low**	**High**
Outcomes							
SanQoL - PL	0.49	0.46	0.53	n/a	n/a	.	.
SanQoL - ST	0.84	0.79	0.89	-133	355	n/a	n/a
SanQoL - CSB	0.78	0.70	0.86			105	166
Costs (for 7,200 people)							
Annuitised capital cost - PL	2,586	2,068	3,103	n/a	n/a	.	.
Annuitised capital cost - ST	101,111	80,889	121,333	-593	-255	n/a	n/a
Annuitised capital cost - CSB	251,705	201,364	302,046			122	171
Household recurrent financial expenditure - PL	11,079	6,072	16,086	n/a	n/a	.	.
Household recurrent financial expenditure - ST	38,619	32,000	45,239	-482	-365	n/a	n/a
Household recurrent financial expenditure - CSB	64,586	46,797	82,376			136	158
Cleaning VOT (recurrent) - PL	7,614	3,807	11,422	n/a	n/a	n/a	n/a
Cleaning VOT (recurrent) - ST	9,828	4,914	14,741				
Cleaning VOT (recurrent) - CSB	9,997	4,998	14,995				
Sludge treatment cost (annual recurrent) - ST & CSB	1,808	1,175	2,440				
Other parameters							
Discount rate	3%	0.01%	10%	-629	-350	122	215
Useful life - PL	5	7	3	n/a	n/a	.	.
Useful life - ST	15	10	20			n/a	n/a
Useful life - CSB	15	10	20			96	277
annual SanQoL decline	0%	n/a	2%			n/a	205
2015 MZN / USD exchange rate	40.1	32.1	48.1	-413	-439	142	153

E. Reporting against CHEERS

Section/item	Item No.	Recommendation	Reported on page no. / line no.
Title and abstract			
Title	1	Identify the study as an economic evaluation or use more specific terms such as "cost-effectiveness analysis", and describe the interventions compared.	Title
Abstract	2	Provide a structured summary of objectives, perspective, setting, methods (including study design and inputs), results (including base case and uncertainty analyses), and conclusions.	Abstract
Introduction			
Background and objectives	3	Provide an explicit statement of the broader context for the study. Present the study question and its relevance for policy or practice decisions.	p.193
Methods			
Target population and subgroups	4	Describe characteristics of the base case population and subgroups analysed, including why they were chosen.	p.194
Setting and location	5	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.	p.195
Study perspective	6	Describe the perspective of the study and relate this to the costs being evaluated.	p.195
Comparators	7	Describe the interventions or strategies being compared and state why they were chosen.	p.194ff
Time horizon	8	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.	p.195
Discount rate	9	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.	p.199
Choice of outcomes	10	Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.	p.196
Measurement of effectiveness	11a	*Single study-based estimates:* Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data.	p.197

Section/item	Item No.	Recommendation	Reported on page no. / line no.
	11b	*Synthesis-based estimates:* Describe fully the methods used for identification of included studies and synthesis of clinical effectiveness data.	n/a
Measurement and valuation of preference based outcomes	12	If applicable, describe the population and methods used to elicit preferences for outcomes.	p.197
Estimating resources and costs	13a	*Single study-based economic evaluation:* Describe approaches used to estimate resource use associated with the alternative interventions. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	p.198ff
	13b	*Model-based economic evaluation:* Describe approaches and data sources used to estimate resource use associated with model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	n/a
Currency, price date, and conversion	14	Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. Describe methods for converting costs into a common currency base and the exchange rate.	p.199
Choice of model	15	Describe and give reasons for the specific type of decision-analytical model used. Providing a figure to show model structure is strongly recommended.	p.195, supp. mat. B
Assumptions	16	Describe all structural or other assumptions underpinning the decision-analytical model.	p.195ff, supp. mat. B
Analytical methods	17	Describe all analytical methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (such as half cycle corrections) to a model; and methods	p.199

Section/item	Item No.	Recommendation	Reported on page no. / line no.
		for handling population heterogeneity and uncertainty.	
Results			
Study parameters	18	Report the values, ranges, references, and, if used, probability distributions for all parameters. Report reasons or sources for distributions used to represent uncertainty where appropriate. Providing a table to show the input values is strongly recommended.	Table 2
Incremental costs and outcomes	19	For each intervention, report mean values for the main categories of estimated costs and outcomes of interest, as well as mean differences between the comparator groups. If applicable, report incremental cost-effectiveness ratios.	Table 3, Table 4
Characterising uncertainty	20a	*Single study-based economic evaluation:* Describe the effects of sampling uncertainty for the estimated incremental cost and incremental effectiveness parameters, together with the impact of methodological assumptions (such as discount rate, study perspective).	p.205ff
	20b	*Model-based economic evaluation:* Describe the effects on the results of uncertainty for all input parameters, and uncertainty related to the structure of the model and assumptions.	n/a
Characterising heterogeneity	21	If applicable, report differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups of patients with different baseline characteristics or other observed variability in effects that are not reducible by more information.	n/a
Discussion			
Study findings, limitations, generalisability,	22	Summarise key study findings and describe how they support the conclusions reached. Discuss	p.207ff

Section/item	Item No.	Recommendation	Reported on page no. / line no.
and current knowledge		limitations and the generalisability of the findings and how the findings fit with current knowledge.	
Other			
Source of funding	23	Describe how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis. Describe other non-monetary sources of support.	p.210
Conflicts of interest	24	Describe any potential for conflict of interest of study contributors in accordance with journal policy. In the absence of a journal policy, we recommend authors comply with International Committee of Medical Journal Editors recommendations.	p.210

Appendix F: Consent form

LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE

WRITTEN INFORMED CONSENT FORM

Study title: Cost-effectiveness of an Urban Sanitation Intervention in Maputo

Principal Investigator: Oliver Cumming

	Participant initials box
1. I confirm that I have been read and understand the participant information sheet dated May '18 (Version 2). I have had the opportunity to consider the information, ask questions and have had these answered fully.	
2. I understand that my participation is voluntary and I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.	
3. I agree to take part in the above study	

__________________ **Participant's name**   __________________ **Participant's signature**   __________________ **Date**

By signing and dating below, I, the interview, confirm that the participant has read the participant information sheet and has agreed with the three statements listed above.

__________________ **Name of interviewer**   __________________ **Interviewer's signature**   __________________ **Date**

1 copy for participant, 1 copy for Principal Investigator

Bairro ID: ____________ Quarteiro no. __________

Compound ID: __________ Household ID. __________

Appendix G: Quantitative survey questionnaires

Surveys were undertaken on the mWater surveyor app. Skipping works on basis of “if” conditions rather than “skip to” functions – these are removed for clarity.

SanQoL 2019 survey

Section 1: introductory questions (C)

C1: Enumerator name

- Carla
- Euclímia
- JP
- Faustino

C1a: Time at start of interview

C2: Bairro name

- Aeroporto A
- Aeroporto B
- Chamanculo A
- Chamanculo B
- Chamanculo C
- Chamanculo D
- Malanga
- Maxaquene A
- Maxaquene B
- Maxaquene C
- Maxaquene D
- Minkadjuine
- Munhuana
- Unidade 7
- Urbanização
- Xipamanine

C2a: Compound code according to database list

C3: Compound type, according to database list

- Control
- Intervention

C3a: Did you already inspect the bathroom on this compound?

- No
- Yes

C3b: Has the intervention/control status of the bathroom already been verified?

- No
- Yes

C3c: Take a photo of the bathroom, capturing the floor and walls

C3d: Does the intervention / control status match the database list? (nb. water seal / sifão criterion)

- No – end interview
- Yes

C4: Take the coordinates of this location

C6: Participant code

C7: Has the participant consented?
- No
- Yes

Section 2: demographic questions (Q)

Q1: Gender of participant
- Male
- Female

Q1a: What is your name?

Q1b: What is your mobile number?

Q2: How old are you?

Q3: How many years have you lived on this compound?

Q4: What is your marital status?
- Single, never married
- Married
- Living maritally
- Separated
- Widowed
- Divorced

Q5: How many children do you have less than 14 years of age?

Q6: How many people live in your household? (eat meals together)

Q8: What level of schooling have you attained?
- None
- Primary (incomplete)
- Primary (complete)
- Secondary (incomplete)
- Secondary (complete)
- Technical qualification (incomplete)
- Technical qualification (complete)
- Higher education (incomplete)
- Higher education (complete)

Q9: What is your relationship with most other households on this compound?
- Blood relatives
- Relatives through marriage
- Unrelated neighbours
- No other households

Q10: How many households live on this compound?

Q11: How many people in total live in this compound?

Q11a: How many men aged 18+ have lived on this compound for 4+ years?

Q11b: How many women aged 18+ have lived on this compound for 4+ years?

Q13: Do you own or rent your house?
- Homeowner
- Rented from APIE
- Rented from private landlord
- Living rent-free

Q14: Does the landlord live on this compound?
- No
- Yes

Section 3: assets and water supply (A)

A1: Does your household have...?
- Electricity
- Radio
- Television
- Mobile phone
- Non-mobile phone
- Fridge or freezer
- None of the above

A2: Does any member of this household own?
- Watch
- Bicycle
- Motorbike
- Car or truck
- None of the above

A5: Do you cook indoors, in a separate house or outside
- Indoors
- Indoors in another house
- Outdoors

A6: Do you have a separate room that serves as a kitchen?
- No
- Yes

A7: What is the primary material your floor is made from?
- Earth / uncovered
- Rudimentary Wood
- Parquet / treated wooden boards
- Cement / stone / tiles
- Carpet / mats
- Other (please specify)

A8: What is the primary material your exterior walls are made from?
- Straw / sticks / bamboo / palm
- Wood / zinc sheet / adobe blocks

- Wattle and daub
- Concrete blocks / fired bricks
- Other (please specify)

A9: What is the primary material your roof is made from?
- Weaving of grass or leaves
- Sheets of zinc or lusalite
- Tiles
- Concrete
- Other (please specify)

A10: What is the main source of water used by your household for *drinking*?
- Tap in the dwelling
- Tap in the yard
- Neighbour's tap
- Public tap / standpost
- Borehole
- Protected well / spring
- Rainwater
- Packaged water (bottles / sachets)
- Delivered water (trucks / small carts)
- Unprotected well or spring
- Surface water (river, pond)

A14: How many *hours per day* was water usually available from this source in the past 30 days?
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- more than 8

Section 4: observations (O)

O1: What kind of toilet facility do members of your household usually use?
- WSUP toilet
- Flush toilet with water seal
- Pit latrine (or other technology without water seal)
- No facility

O2: Please may I look at this toilet?
- No
- Yes

O2a: Take a photo, capturing the floor and walls

O1a: [Observe] What type of pan or slab does it have?
- Seat pan with water seal
- Squat pan with water seal
- Concrete slab without water seal
- Wood/soil/other slab without water seal

O1b: [Observe] How is the drop hole covered?
- Fitted lid
- Piece of metal, plastic or similar
- No cover on hole
- Other (please specify)

O5: [Observe] What material is the latrine floor made of?
- Soil
- Concrete slab
- Wood
- Tiling
- Other (please specify)

O8: [Observe] What is the condition of the slab / floor?
- In general in good condition
- Cracked or broken slab/floor but no holes in the ground or apparent risk of collapse
- Cracked or broken slab/floor, holes in the ground but no risk of collapse
- Slab/floor appears to be at risk of collapse
- In general in good condition, but without slab of concrete, the floor is sand

O6: [Observe] What material are the toilet walls mainly built from?
- Masonry / concrete blocks
- Complete zinc sheets
- Many pieces of scrap metal
- Reeds / grass / bamboo / palm
- Plastic, bags or cloth
- No walls
- Other (please specify)

O7: [Observe] What material is the toilet roof mainly built from?
- Tiles
- Concrete / cement sheets
- Complete zinc sheets
- Many pieces of scrap metal
- Reeds / grass / bamboo / palm
- Plastic, bags or cloth
- No roof
- Other (please specify)

O9: [Observe] Is water available near the latrine/toilet?
- Water is available in a container
- Water is available from a tap next to the latrine
- Water is not available or container is empty

O10: [Observe] What is the hygienic condition of this latrine? (select all that apply)
- Presence of dirty water
- Presence of solid waste
- Presence of urine
- Presence of used anal cleaning materials
- Presence of feces
- Strong smell of urine
- Strong smell of faeces
- None of the above

Section 5: toilet usually used (G)

G1: Do you share this toilet facility with other households?
- No
- Yes

G2: Including your own household, how many households use this toilet facility?

G3: Including your own household members, how many people use this toilet facility?

G4: Where is this toilet facility located?
- In own dwelling
- In own compound
- Elsewhere

G5: Can anyone who is not a member of your compound use the toilet without asking?
- No
- Yes

G6: Can this toilet be locked from the inside?
- No
- Yes

G7: Can this toilet be locked from the outside?
- No
- Yes

G8: Are there ever times when you need to use the toilet/latrine but you don't have a key?
- No
- Yes

G10: In what year did the residents of this compound begin to use this toilet?
Put year (e.g. 2014) not number of years
- Don't Know

G12: When was the last time the pit/tank of this bathroom was emptied?
- 0-6 months ago
- 7-12 months ago
- 1-2 years ago
- 2-3 years ago
- 3-4 years ago
- 5-6 years ago
- >6 years ago
- Don't Know
- Not Applicable

G13a: Do users clean the toilet?
- No
- Sometimes
- Yes

G13: Is there a rota/schedule for cleaning the latrine?
- No
- Yes, written down
- Yes, but not written down
- Don't Know

G14: Do people adhere to the cleaning schedule/rota?
- Nobody adheres
- Some people adhere
- Everyone adheres
- Don't Know

G15: Is there another toilet that members of your household sometimes use when at home?
- No
- WSUP toilet
- Flush toilet with water seal
- Pit latrine (or other technology without water seal)
- No facility

G15a: What type of pan or slab does it have?
- Seat pan with water seal
- Squat pan with water seal
- Concrete slab without water seal
- Wood/soil/other slab without water seal

G18: What do you normally do with grey water?
- Dump on the ground inside the compound
- Dump on the ground in the street
- Dumps into a drain or soak pit
- Dump in the latrine or toilet
- Use to water plants or trees within compound
- Other (please specify)

Section 6: SanQoL questions

Read to respondent: "Let's talk about your experiences while carrying out your sanitation practices in the past 4 weeks. By this I mean any practices you carry out in the bathroom you usually use when at home. I am interested in how often you experienced the things in the questions in the past 4 weeks. So, please respond with: always, sometimes, rarely or never."

D3: How often do you see other people's faeces when using the bathroom?
negative framing - "always" is bad
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

D5: Can you use the bathroom without smelling other people's faeces?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

D1: Can you use the bathroom without feeling disgust?
- Always
- Sometimes
- Rarely
- Never

- ---Prefer not to answer

D2: How often do you find the bathroom clean when entering?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

D4: Can you feel clean while using this bathroom?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

D7: Can you use the bathroom without coming into contact with faeces?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

D8: Can you use the bathroom without flies causing you disgust?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

D11: How often do you see flies landing on food you are preparing or eating?
negative framing - "always" is bad

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

H1: Can you use the bathroom without worrying that it spreads diseases?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

H5: Can you use the bathroom without negative consequences for your health and wellbeing?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

H4: How often do you worry that your family might catch diseases due to poor sanitation in this neighbourhood in general?
negative framing - "always" is bad

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

P1: Can you use the bathroom in private, without being seen?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

P2: While you are using the bathroom, how often do you worry about people watching you?
negative framing - "always" is bad
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

P3: Can you use the bathroom without people interrupting you?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

S1: Can you use the bathroom without feeling ashamed for any reason?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

S2: Are you confident that the bathroom doesn't reduce neighbours' respect for you?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

S3: How often do you worry about neighbours being able to smell the bathroom?
negative framing - "always" is bad
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

S5: When visitors come, do you ever feel embarrassed providing this bathroom for them to use?
negative framing - "always" is bad
- Always
- Sometimes
- Rarely

- Never
- ---Prefer not to answer

S6: Can you feel proud of your bathroom?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

S7: How often do you worry that using this bathroom is not dignified?
negative framing - "always" is bad
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

T1: Are you able to feel safe while using the bathroom?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

T2: Are you able to feel safe while using the bathroom at night?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

T10: How often do you use a potty or bucket at night for fear of using the bathroom?
negative framing - "always" is bad
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

T5: Are you able to use the bathroom in the daytime without fearing assault?
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

T6: How often do you worry that someone in your family will be assaulted while using the bathroom?
negative framing - "always" is bad
- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

T7: Are you able to use the bathroom without being harassed?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

T9: Are you confident that nobody in your family will be injured by an accident in the bathroom?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

E1: Can you easily ensure that the environment immediately around your house is clean?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

E2: How often are the streets near your house contaminated with human or animal faeces?
negative framing - "always" is bad

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

E3: How often does dirty water from other people's bathrooms contaminate the streets?
negative framing - "always" is bad

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

E5: When it rains hard, how often do you worry that water will flood inside your house?
negative framing - "always" is bad

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

W1: Can you use as much water as you need when using the bathroom?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

M1: Can you easily ensure the bathroom is kept clean?

- Always
- Sometimes

- Rarely
- Never
- ---Prefer not to answer

M3: How often do you have arguments about cleaning the bathroom?
negative framing - "always" is bad

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

M5: Can you use the bathroom whenever you want, without having to wait for other people?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

M6: How often do you have arguments about people wanting to use the bathroom at the same time?
negative framing - "always" is bad

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

X2: Can you use the bathroom independently, without help from others?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

X3: Can you use the bathroom without experiencing physical pain?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

X4: How often do you feel comfortable when using the bathroom?

- Always
- Sometimes
- Rarely
- Never
- ---Prefer not to answer

B1: Have you had diarrhoea in the past 7 days? Diarrhoea is 3 or more liquid stools in a day

- No
- Yes

B2: Have you had a stomach ache in the past 7 days?

- No
- Yes

B3: Have you vomited in the past 7 days?

- No
- Yes

B4: Do you know anybody personally who has been sexually assaulted while using a bathroom in this neighbourhood?
nb. ever, not only in the past 4 weeks

- No
- Yes
- ---Prefer not to answer

B5: Do you know anybody personally who has been physically assaulted while using a bathroom in this neighbourhood?
nb. ever, not only in the past 4 weeks

- No
- Yes
- ---Prefer not to answer

B6: Do you know anybody personally who has been harassed while using a bathroom in this neighbourhood?
nb. ever, not only in the past 4 weeks

- No
- Yes
- ---Prefer not to answer

Section 7: WHO-5 wellbeing (F)

All WHO-5 questions asked with respect to the last 2 weeks. Explain that this set of questions is about frequency, with 6 options.

F1: In the last 2 weeks have you felt cheerful and in good spirits?

- At no time
- Some of the time
- Less than half of the time
- More than half of the time
- Most of the time
- All of the time
- ---Prefer not to answer

F2: ...felt calm and relaxed?

- At no time
- Some of the time
- Less than half of the time
- More than half of the time
- Most of the time
- All of the time
- ---Prefer not to answer

F3: ...felt active and vigorous?

- At no time
- Some of the time
- Less than half of the time
- More than half of the time
- Most of the time
- All of the time

- ---Prefer not to answer

F4: ...woken up feeling fresh and rested?

- At no time
- Some of the time
- Less than half of the time
- More than half of the time
- Most of the time
- All of the time
- ---Prefer not to answer

F5: ...had a daily life filled with things that interest you?

- At no time
- Some of the time
- Less than half of the time
- More than half of the time
- Most of the time
- All of the time
- ---Prefer not to answer

Section 8: EQ-5D health measure (EQ)

Read out: "please select the option that best describes your health TODAY". Explain that these response options are different and there are 5.

EQ1: Do you have problems walking about?

- No problems
- Slight problems
- Moderate problems
- Severe problems
- Unable to do so
- ---Prefer not to answer

EQ2: Do you have problems washing or dressing yourself?

- No problems
- Slight problems
- Moderate problems
- Severe problems
- Unable to do so
- ---Prefer not to answer

EQ3: Do you have problems undertaking your usual activities (e.g. work, study, housework, family or leisure activities)?

- No problems
- Slight problems
- Moderate problems
- Severe problems
- Unable to do so
- ---Prefer not to answer

EQ4: Do you have pain or discomfort?

- no pain or discomfort
- slight pain or discomfort
- moderate pain or discomfort

- severe pain or discomfort
- extreme pain or discomfort
- ---Prefer not to answer

EQ5: Are you anxious or depressed?
- not anxious or depressed
- slightly anxious or depressed
- moderately anxious or depressed
- severely anxious or depressed
- extremely anxious or depressed
- ---Prefer not to answer

Section 9: valuation (V)

[VAS instructions to enumerator on printed sheet (see Appendix D: supplementary information to Chapter 7)]

V1: Input score from printed-out 'visual analogue scale' with smiley faces

Read out: "Now let's do an exercise to understand the value you place on different attributes of a bathroom. Don't think about the bathroom you have now, but bathrooms in general. This is board has 1 to 10 marked on it. 10 is 'most important for a good bathroom' and 1 is 'least important for a good bathroom'. Here you have seven attributes. I would like to choose the most important from the seven, and put it at position 10. Then choose the least important and we'll put at position 1. Then we'll put the others in between. You may have as many attributes on one position as you like. However, there should be at least one attribute at position 10 and at least one attribute at position 1."

V2: Velcro scale - attribute listed top
- Disgust - no flies or smells from the bathroom
- Health - the bathroom doesn't spread diseases
- Shame - I don't feel embarrassed because of the bathroom
- Cleanliness - I feel clean using the bathroom
- Safety - I feel safe using the bathroom
- Privacy - Nobody can see me in the bathroom
- Respect - I don't feel disrespected because of the bathroom
- Not Applicable

V3: Input top attribute position on scale

V4: Attribute listed 2nd
- Disgust - no flies or smells from the bathroom
- Health - the bathroom doesn't spread diseases
- Shame - I don't feel embarrassed because of the bathroom
- Cleanliness - I feel clean using the bathroom
- Safety - I feel safe using the bathroom
- Privacy - Nobody can see me in the bathroom
- Respect - I don't feel disrespected because of the bathroom
- Not Applicable

V5: Input 2nd attribute position on scale

V6: Attribute listed 3rd
- Disgust - no flies or smells from the bathroom
- Health - the bathroom doesn't spread diseases
- Shame - I don't feel embarrassed because of the bathroom
- Cleanliness - I feel clean using the bathroom

- Safety - I feel safe using the bathroom
- Privacy - Nobody can see me in the bathroom
- Respect - I don't feel disrespected because of the bathroom
- Not Applicable

V7: Input 3rd attribute position on scale

V8: Attribute listed 4th
- Disgust - no flies or smells from the bathroom
- Health - the bathroom doesn't spread diseases
- Shame - I don't feel embarrassed because of the bathroom
- Cleanliness - I feel clean using the bathroom
- Safety - I feel safe using the bathroom
- Privacy - Nobody can see me in the bathroom
- Respect - I don't feel disrespected because of the bathroom
- Not Applicable

V9: Input 4th attribute position on scale

V10: Attribute listed 5th
- Disgust - no flies or smells from the bathroom
- Health - the bathroom doesn't spread diseases
- Shame - I don't feel embarrassed because of the bathroom
- Cleanliness - I feel clean using the bathroom
- Safety - I feel safe using the bathroom
- Privacy - Nobody can see me in the bathroom
- Respect - I don't feel disrespected because of the bathroom
- Not Applicable

V11: Input 5th attribute position on scale

V12: Attribute listed 6th
- Disgust - no flies or smells from the bathroom
- Health - the bathroom doesn't spread diseases
- Shame - I don't feel embarrassed because of the bathroom
- Cleanliness - I feel clean using the bathroom
- Safety - I feel safe using the bathroom
- Privacy - Nobody can see me in the bathroom
- Respect - I don't feel disrespected because of the bathroom
- Not Applicable

V13: Input 6th attribute position on scale

V14: Attribute listed 7th
- Disgust - no flies or smells from the bathroom
- Health - the bathroom doesn't spread diseases
- Shame - I don't feel embarrassed because of the bathroom
- Cleanliness - I feel clean using the bathroom
- Safety - I feel safe using the bathroom
- Privacy - Nobody can see me in the bathroom
- Respect - I don't feel disrespected because of the bathroom
- Not Applicable

V15: Input 7th attribute position on scale

V16a: Take a photo of the completed board

V16: [question for interviewer - do not read] Do you think the respondent understood the velcro scaling exercise?

- Completely understood
- Mostly understood
- Partially understood
- Not at all understood
- Not Applicable

Section 10. End

Z0: Enumerator free text comments on whole interview

Z1: Time at end of interview

Costing modules within 2018 MapSan process evaluation survey

Below are the questions relevant to the cost-effectiveness study from a longer survey undertaken for the process evaluation of MapSan.

Section 1: costs (C)

C01: Were you living on this compound when the toilet/ latrine you use most often was constructed?

C01a: What happened to the latrine you used before?

- Abandoned
- Covered with earth
- Covered with garbage
- Covered with other
- Emptied
- Other

C02: When the latrine/ toilet you use most frequently was constructed how much was paid in total by everybody who contributed?

C03: How much money did your HH originally contribute for construction of the latrine/ toilet you use most often?

C03a: Did you HH contribute labour or materials for construction of the latrine you use most often?

C04: How much money did the landlord originally contribute for construction of the latrine you use most often?

C05: How much money did other HH contribute on average for construction of the latrine/toilet you use most often?

- More money than your household
- About the same amount as your household
- Less money than your household
- No money was contributed

C06: Did the need to contribute to the cost of building the latrine cause any delay to construction of the latrine?

- No
- Yes, it was difficult for this HH to contribute to the construction of the latrine
- Yes, it was difficult for other HHs to contribute to the construction of the latrine
- Yes it was difficult for the landlord to contribute to the cost of the latrine

Section 2: emptying (F)

F01: In what year did the residents of this compound begin to use the latrine/ toilet you use most often?

F01a: In what part of the year?

- Beginning
- Middle
- End

F02: How many times has a new septic tank been constructed in this compound in the last 3 years

F03: How many times has a latrine/septic tank on this compound been emptied in the last three years?

F04: When was the last time the latrine/septic tank that you use most often was emptied?

- Never
- With the last week
- Within the last month
- Within the last six months
- Within the last year
- Within the last two years
- Within the last 5 years
- More than five years ago
- Pit/ tank was replaced not emptied

F04a: Which of these services exist in your barrio?

- Manual emptying with sacks
- Manual pumping
- Mechanical pumping
- Vacuum truck pumping

F05: last time, who emptied the septic tank you use most frequently?

- HH member
- Informal emptier
- Formal emptier
- Other

F06: Last time the latrine/septic tank you use most frequently was emptied what equipment was used to empty it?

- By hand, using sacks, buckets or similar hand tools
- By hand, using a hand pump (such as a gulper)
- Mechanically, using a small mechanical pump (such as a trash pump)
- Mechanically, using a vacuum truck
- Unsure

F06a: Why did you not choose a formal emptying service the last time you emptied the tank

- Cost
- Service not available near my house
- Other

F08: Last time the pit or septic tank you use most often was emptied, where was the faecal waste disposed?

- Buried inside the compound
- Buried nearby, but not inside the compound
- Dumped on the ground nearby the compound
- Taken to treatment plant
- Taken outside the compound (destination unknown)
- Taken outside the compound (specify destination)
- Uncertain

F10: Last time the pit/tank of latrine/ toilet you use most often was emptied, how much was paid in total?

F11: The last time the pit/tank of the latrine/toilet you use most often on this compound was emptied, how much was contributed by *this household*?

F12: The last time the pit/tank of the latrine/toilet you use most often on this compound was emptied, how much was contributed by *the landlord*?

- More money than your household
- About the same amount as your household

- Less money than your household
- No money was contributed

F13: The last time the pit/tank of the latrine/toilet you use most often on this compound was emptied, how much was contributed by *each other household, on average*?

- More money than your household
- About the same amount as your household
- Less money than your household
- No money was contributed

F14: How was payment made last time?

- Fixed price
- Cost per volume removed

F15: Was payment made in instalments?

- No, paid in one go
- Yes, in 2 parts
- Yes in 3 parts
- Yes, in more than 3 parts

F17: The next time you need to empty a pit or septic tank on this compound, what specific action do you expect to be performed?

- Close and open another pit in the compound
- Close and use neighbour's latrine
- Latrine/ Septic tank will be emptied by a HH member
- Latrine/ septic tank will be emptied by informal emptier
- Latrine/ Septic tank will be emptied by formal emptier
- Latrine/ septic tank will be replaced not emptied
- Other"

Section 3: Cleaning and maintenance (CL)

CL08: In a week, how much time does your household spend on average on cleaning and maintaining the toilet/latrine you use most often?

CL09: In a week, how much time does your landlord spend on average on cleaning and maintaining the toilet/latrine you use most often?

CL10: In a week, how much time do the other HHs spend on average on cleaning and maintaining the toilet/latrine you use most often?

CL11: How many times in the last week did you personally clean the latrine/toilet?

CL12: Has your HH spent money to buy cleaning products for the latrine? (multicheck item)

- Omo/detergent
- Broom
- Handmade broom
- Bleach
- Bucket
- Mop
- No money is spent on cleaning items
- Other

CL13: How much money do you spend on X every time you buy it? (asked per multicheck item in CL12)

CL14: How often do you buy this item? (asked per multicheck item in CL12)

- Weekly
- Monthly
- Twice a year
- Once a year
- Less than once a year

CL15: Do others contribute to this cost? (asked per multicheck item in CL12)

- This household pays the total
- The cost is divided equally between families
- This household paid more than others
- Other households paid more than us

CL16 – CL33: equivalent of CL13-CL15 per multicheck item in CL12

CL34: In general, who makes decisions about repairs to the toilet/latrine you use most often?

- Landlord
- Compound chief
- Men on the compound
- Women on the compound
- Everyone on the compound

CL35: Do you make a regular contribution to a toilet/latrine maintenance fund?

- No
- Yes

CL36: How much does your HH contribute monthly?

CL37: Do all other households make the same contribution?

- No
- Yes
- Everyone except landlord

CL38: [asked to those with a fund] From the maintenance fund, which of the following items were repaired in the last year?

- Door
- Windows
- Floor
- Walls
- Water storage/connection
- Water tank
- Slab / pedestal
- Lock
- No money was spent on repairs
- Other

CL39: [asked to those without a fund] In the last year, has your family spent to repair or replace the following items on the latrine/toilet which you use most often? (multicheck)

- Door
- Windows
- Floor
- Walls
- Water storage/connection
- Water tank
- Slab / pedestal
- Lock

- No money was spent on repairs
- Other

CL40: How much money did your HH contribute for X? (asked per multicheck item in CL39)

CL41: How much money did other HHs contribute? (asked per multicheck item in CL12)
- This household paid the total
- The cost was divided equally between families
- This household paid more than others
- Other households paid more than us

CL42 – CL53: equivalent of CL40-CL41 per multicheck item in CL39

CL54: How much does the landlord contribute to maintaining the latrine/toilet used most often?
- More money than this household
- About the same as this household
- Less money than this household
- Landlord does not spend money on this