


Evaluations of water, sanitation and hygiene interventions should not use diarrhoea as (primary) outcome

Samuel I Watson,¹ Ryan T T Rego,² Timothy Hofer,³ Richard J Lilford ¹

To cite: Watson SI, Rego RTT, Hofer T, *et al*. Evaluations of water, sanitation and hygiene interventions should not use diarrhoea as (primary) outcome. *BMJ Global Health* 2022;**7**:e008521. doi:10.1136/bmjgh-2022-008521

Handling editor Seye Abimbola

Received 14 January 2022

Accepted 19 April 2022

ABSTRACT

Water, sanitation and hygiene interventions have been the subject of cluster trials of unprecedented size, scale and cost in recent years. However, the question ‘what works in water, sanitation, hygiene (WASH)?’ remains poorly understood. Evaluations of community interventions to prevent infectious disease typically use lab-confirmed infection as a primary outcome; however, WASH trials mostly use reported diarrhoea. While diarrhoea is a significant source of morbidity, it is subjected to significant misclassification error with respect to enteric infection due to the existence of non-infectious diarrhoea and asymptomatic infection. We show how this may lead to bias of estimated effects of interventions from WASH trials towards no effect. The problem is further compounded by other biases in the measurement process. Alongside testing for infection of the gut, an examination of the causal assumptions underlying WASH interventions present several other reliable alternative and complementary measurements and outcomes. Contemporary guidance on the evaluation of complex interventions requires researchers to take a broad view of the causal effects of an intervention across a system. Reported diarrhoea can fail to even be a reliable measure of changes to gastrointestinal health and so should not be used as a primary outcome if we are to progress our knowledge of what works in WASH.

SUMMARY BOX

- ⇒ Despite several recent cluster trials of unprecedented size, scale and cost evaluating water, sanitation, hygiene (WASH) interventions, the question ‘what works in WASH?’ remains poorly understood.
- ⇒ Evaluations of community interventions to prevent infectious disease typically use lab-confirmed infection as a primary outcome; however, WASH trials mostly use reported diarrhoea.
- ⇒ Diarrhoea is a significant source of morbidity, but it is subjected to significant misclassification error with respect to enteric infection due to the existence of non-infectious diarrhoea and asymptomatic infection.
- ⇒ We show how misclassification of diarrhoea leads to a bias of estimated effects of interventions from WASH trials towards no effect, which is compounded by further biases in the measurement process.
- ⇒ Reported diarrhoea can fail to be a reliable measure of changes to gastrointestinal health and so should not be used as a primary outcome if we are to progress our knowledge of what works in WASH.



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Institute of Applied Health Research, University of Birmingham, Birmingham, UK

²Center for Global Health Equity, University of Michigan, Ann Arbor, Michigan, USA

³Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, Michigan, USA

Correspondence to

Dr Samuel I Watson;
s.i.watson@bham.ac.uk

INTRODUCTION

Diarrhoeal disease remains one of the most prolific killers of children under 5.^{1 2} The predominant strategy to prevent these deaths is to improve water, sanitation and hygiene (WASH) infrastructure and related behaviours.³ Significant efforts are focused in low and middle-income countries (LMICs) where the burden of disease is highest. However, despite huge investment into intervention development and evaluation, the answer to the question ‘what works in WASH?’ remains poorly understood.^{4 5}

In the last few years, three cluster trials of unprecedented size, scale and cost—WASH-Benefits Bangladesh,⁶ WASH-Benefits Kenya⁷ and the Sanitation Hygiene Infant

Nutrition Efficacy (SHINE) trials⁸—found little evidence of benefit from any of the WASH components of the different interventions. Some commentaries have suggested that the interventions, which are among the most common types of WASH interventions in LMIC settings,³ were of inadequate intensity, poorly tailored to the local modes of disease transmission and possibly not acceptable to the targeted community.^{4 5} While we would agree with these conclusions in general, in this article, we will argue that there is an additional problem with these and many other trials of WASH interventions: the use of reported diarrhoea as the primary outcome measurement to assess effectiveness.

DIARRHOEA AS AN OUTCOME

Diarrhoea is typically defined as three or more loose or watery stools in 24 hours period and it has both infectious and non-infectious causes.⁹ There are several different

methods used to measure episodes of diarrhoea among the under-fives. By far, the most common is retrospective self-report in which caregivers or family members are asked to recount whether an infant has had diarrhoea in the preceding period (typically between 24 hours and 2 weeks).¹⁰ The Demographic and Health Surveys, UNICEF’s Multiple Indicator Cluster Surveys¹¹ and the three large trials mentioned above use this community survey method. Other approaches include direct observation of collected stool samples by field workers,¹² prospective diary-based methods¹⁰ or recording hospitalisation rates for diarrhoeal disease at local health centres and hospitals.^{13–15}

WASH interventions are designed to interrupt faecal-oral transmission of pathogens in order to prevent enteric infection and subsequent diarrhoeal disease. So the typical objective of a WASH trial is to demonstrate that an intervention causes a reduction in diarrhoea rates. However, the only type of diarrhoea that can *directly* be reduced by WASH interventions is infectious diarrhoea. Our contention is that measuring all episodes of diarrhoea instead of infectious diarrhoea or enteric infection may result in significant measurement error, which limits its usefulness for assessing WASH interventions.

THE PROBLEM WITH DIARRHOEA

Diarrhoea is the primary clinical presentation of symptomatic diarrhoeal disease and a direct cause of significant morbidity through dehydration and malnutrition. So it might seem like an obvious outcome to assess changes in diarrhoeal disease rates. However, this same logic of measuring an outcome with multiple causes does not apply to trials in other disease areas. For example, randomised trials of interventions to reduce transmission of COVID-19 do not use cough or febrile illness as an outcome, nor do interventions targeted at reducing human papillomavirus examine vaginal bleeding as an outcome; they use laboratory confirmed outcomes, including PCR-based methods^{16 17}. To see why, we can think of symptom reporting as a form of diagnostic test.

Consider diarrhoea and enteric infection by pathogenic organisms: some people are carriers so not all cases of enteric infection present with diarrhoea (the ‘sensitivity’ is below 100%), and some people have diarrhoea in the absence of enteric infection (the ‘specificity’ is below

100%). **Table 1** shows an example 2x2 contingency table. If we use the prevalence (or rate) of diarrhoea to try to estimate the prevalence (or rate) of enteric infection, our estimate will be biased. If *p* is the true prevalence of infection and \hat{p} is our estimate of the prevalence based on diarrhoea, then:

$$\hat{p} = Se \times p + (1 - Sp) \times (1 - p),$$

where *Se* is the sensitivity and *Sp* the specificity and both are between 0 and 1. Any measure based on this biased estimator, such as rate or risk ratios, will itself be biased. For example, if *p*₀ and *p*₁ are the true prevalences in the control and treatment arms of a trial, respectively, then:

$$\frac{\hat{p}_1}{\hat{p}_0} = \frac{Se \times p_1 + (1 - Sp) \times (1 - p_1)}{Se \times p_0 + (1 - Sp) \times (1 - p_0)}$$

which is not equal to *p*₁/*p*₀, the true relative risk, unless the specificity is 100%. Indeed, the worse the sensitivity and specificity are, the more the ratio will be biased towards 1, that is, no effect.

The problem described here is often referred to as classification error, which is the more specific description for measurement error that occurs when using dichotomous or categorical measurements.¹⁸ Neuhaus¹⁹ demonstrated that in logistic regression misclassification in the response variable introduces both a loss of efficiency and a highly biased attenuation of the effect estimators.¹⁹ This bias will be present when estimating any benefit of WASH interventions when diarrhoea due to infection outcome is misclassified.

EVIDENCE ON THE ‘DIAGNOSTIC PERFORMANCE’ OF DIARRHOEA

The problem goes away if the sensitivity and specificity of diarrhoea are close to 100%. Unfortunately, these values are unknown as they likely vary between countries and contexts as well as between methods of ascertaining diarrhoea rates. The lack of certainty alone should generate caution when using these outcomes. However, recent evidence suggests that they are very unlikely to be close to 100%. For example, the two-by-two contingency table above can also be used to derive an OR (AD/BC in **table 1**): if sensitivity and specificity were 100%, we would expect the ORs to tend to infinity, that is, be very large as B and C would be 0. Relatively small ORs, therefore, indicate poor ‘diagnostic performance’. A recent systematic review of case control studies comparing enteropathogen presence in mostly hospitalised cases of diarrhoea versus controls without diarrhoea found ORs for different pathogens to predominantly fall in the range of 0.5–5.0,²⁰ providing evidence of misclassification bias that differs by pathogen. Only very aggressive pathogens like *Vibrio cholerae* had very large ORs (around 50). In a recent study we conducted in the Cox’s Bazar camps for Forcibly Displaced Rohingya Population for Myanmar in Bangladesh, we estimated the all-pathogen sensitivity and specificity of carer-reported diarrhoea to be 0.49 (95% CI 0.39 to 0.66) and 0.65 (0.41 to 0.85), respectively.¹² The

Table 1 A two-by-two contingency table for enteric infection and diarrhoea

		Diarrhoea	
		Yes	No
Enteric infection	Yes	Symptomatic infection (A)	Asymptomatic infection (B)
	No	Non-infectious diarrhoea (C)	Healthy (D)
Sensitivity is A/(A+B) and specificity is D/(C+D).			

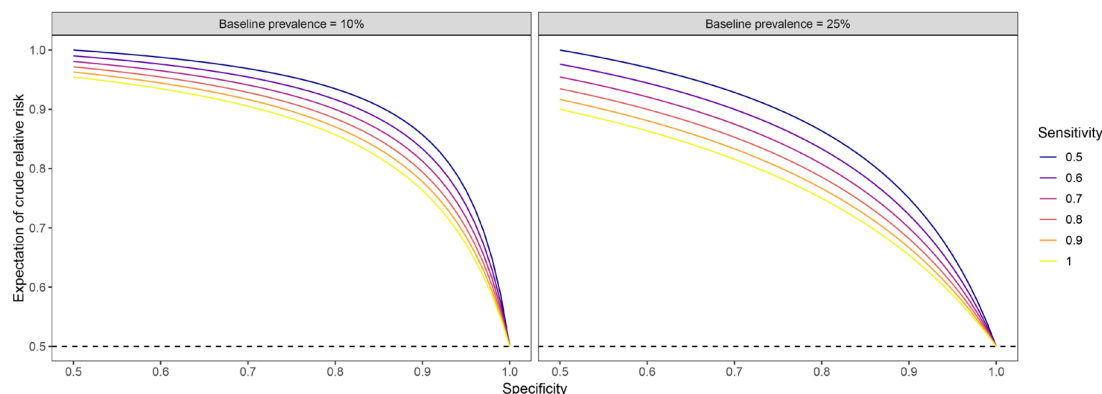


Figure 1 Relationship between baseline prevalence, sensitivity and specificity and the estimated relative risk. In all cases, the true relative risk is 0.5.

incidence of diarrhoea was unusually high in this setting and if sensitivity and specificity vary with prevalence, then these results might underestimate sensitivity and specificity in other settings.

The diarrhoea prevalences reported in the three large trials^{6–8} were 5% to 10%, which puts a requisite lower limit on the specificity of 90% to 95%. However, even with these more optimistic figures, we might still suspect quite significant bias. Figure 1 shows a hypothetical example in which the true relative risk of enteric infection between a treatment (eg, water and sanitation improvement) and a control group is 0.5 and the baseline prevalence of infection is either 10% or 25%. We show how different values of sensitivity and specificity of a diarrhoea outcome affect the estimated relative risk of the study. In the case where the baseline is 10%, even if sensitivity and specificity are as high as 90%, the estimated relative risk is attenuated from 0.5 to 0.78. The same effect would be apparent for relative risks greater than 1 as well.

The problems above are further compounded when using self-reported and survey-based measures of diarrhoea (due to additional measurement error introduced by the difference between ‘diarrhoea’ and ‘reported diarrhoea’ in figure 2). Our aforementioned study compared agreement statistics (Cohen’s *d*) for, among other methods, a standard retrospective recall survey, a survey

augmented with pictorial representations of different stools and visual inspection by trained field workers; we estimated values of between -0.1 and 0.1 , indicating very poor agreement.¹² Changes in the length of recall period or the frequency of questioning can also affect estimated rates of diarrhoea.^{21–23} Therefore, it is very likely that the above three WASH trials, and other comparable studies, have underestimated any intervention effect. Alternatives to survey-based diarrhoea assessment also have their own issues. For example, hospital-reported rates are low, occurring about once for every 50 cases of carer-reported diarrhoea, suggesting that this measurement suffers from severe underascertainment of community cases and/or under-reporting by hospitals.¹⁰ The biases we describe exist in addition to others that may affect trials in this area, such as selective attrition or selection bias.

ALTERNATIVES TO DIARRHOEA AS A PRIMARY OUTCOME

The aim of a WASH intervention is to prevent symptomatic disease and morbidity.^{24 25} One might argue that rates of infection with pathogenic bacteria are, therefore, only of instrumental importance, while symptomatic illness is of primary clinical importance and as such the presence of infectious diarrhoea should be the ‘primary outcome’. However, this argument fails on two fronts.

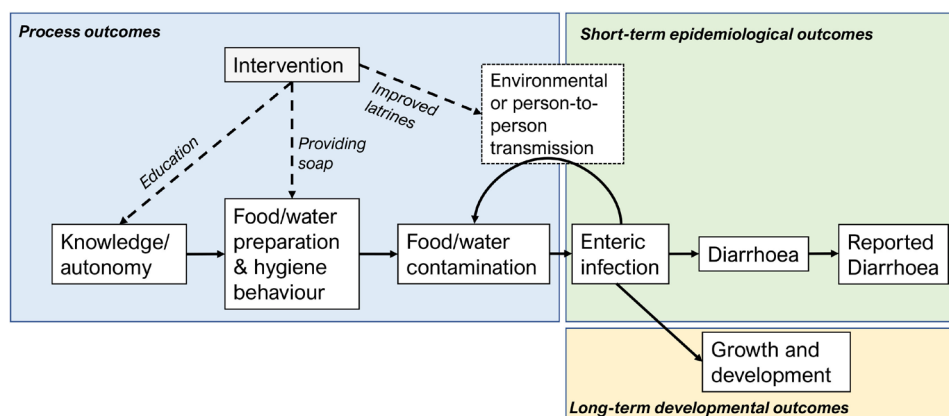


Figure 2 Simplified, illustrative causal diagram linking diarrhoeal disease intervention to outcomes with examples of such interventions.

First, with few exceptions, studies that use reported diarrhoea rates only measure whether the symptom is present and do not confirm the underlying infection status, leading to the misclassification problem described above. Second, newer models of gut health and the microbiome suggest that the presence of enteric pathogens reflects a significant loss of colonisation resistance, which may have clinical significance due to the immunologic and microbiome changes that led to the loss of resistance and which themselves can increase the pathogenic potential of any non-commensal gut resident.^{26 27} Asymptomatic infected people may also be an important reservoir for pathogens that cause symptomatic infections in other people in areas where water and sanitation are inadequate.^{28 29} WASH interventions, by reducing transmission, may well, therefore, significantly reduce the amount of asymptomatic infection. Thus, preventing infection, whether symptomatic or not, has further relevance to lowering morbidity and mortality, so it cannot be claimed to be only of instrumental importance. We should, therefore, consider alternatives.

PROCESS OUTCOMES

Figure 2 shows a simplified and illustrative causal model for the effects of an intervention designed to tackle diarrhoeal disease with some examples of WASH interventions. The first set of outcomes are the immediate, ‘upstream’ effects, such as changing behaviour. In the language of complex interventions, these are often called ‘process outcomes’, but it could also be referred to as an upstream ‘mediating variable’ in line with the burgeoning statistical and epidemiological literature on causal modelling.³⁰ Some cluster trials of specifically behavioural WASH interventions have used these as primary outcomes.^{31 32} The three large WASH trials captured behavioural outcomes generally as measures of adherence to the intervention, and while they show improvement over time in an intervention cluster, they stop short of formally comparing them between intervention and control. The largest effects of the intervention are likely to be seen on these process outcomes³³ and they are relatively inexpensive to collect; however, some assessments of behaviour may be subjective and, thus, subject to similar biases as diarrhoea. Influencing the process outcomes is also only a necessary, but not sufficient, condition for an effect to materialise on the more downstream outcomes.

SHORT-TERM EPIDEMIOLOGICAL AND CLINICAL OUTCOMES

We then have the short-term epidemiological outcomes, such as enteric infection and diarrhoea. Direct assessment of enteric infection is much less common in WASH trials than diarrhoea; however, there are some notable examples.^{34 35} A secondary analysis of the SHINE trial data published in a separate article examined enteric infection captured from stool samples in the trial. They found evidence of reduced prevalence of parasites, but

little evidence of change in viral and bacterial carriage rates.³⁶ Enteric infection presents an attractive option as it is ‘objective’ in the sense of being lab based rather than survey based.

One potential barrier to the use of microbiological outcomes, such as the presence of gut pathogens, is their cost and resource requirements. Stool testing requires the storage and refrigerated shipping of large numbers of samples to a lab equipped with trained staff and expensive equipment. Indeed, in many LMIC settings, such lab facilities are not available at the required scale. One could limit their inferences to the more upstream outcomes in these circumstances or reduce the sample size or number of pathogens to test to reduce costs. Alternatives include sample pooling and environmental surveillance.³⁷ Another alternative may be rapid field tests, including immunochromatographic assays, to establish infection. While these tests also have imperfect sensitivity and specificity, their diagnostic performance can be established in a lab and used to adjust or correct results at the end by using the imperfect model described above.^{18 38} Indeed, we are conducting a pilot study of such a data collection process.

LONG-TERM DEVELOPMENTAL OUTCOMES

Finally, there are the long-term developmental outcomes around physiological and cognitive development of which linear growth is frequently reported.^{5 39} Long-term health and well-being outcomes may be preferred as they are intrinsically valuable, whereas the other outcomes may be considered only of instrumental value. While obviously important, these long-term outcomes result from the confluence of a range of factors. Any effect here is likely to be small and hidden among significant noise. The shorter term epidemiological outcomes might, therefore, represent a good trade-off.

There are evidently many potential outcomes a trial could use, which are often captured. However, most WASH trials use only a single ‘primary’ outcome on which the main conclusion of the trial is based, which may be a consequence more of the requirements of null hypothesis significance testing,⁴⁰ rather than a principled approach to scientific investigation of WASH.

An important corollary to this discussion is that it is difficult to ascertain the effectiveness of a complex intervention, or an intervention in a complex causal path, by looking only at one outcome, especially one with significant measurement error. For example, consider a behavioural change WASH intervention that aims to educate caregivers about improving hygiene and reducing contamination of food and water. If we choose diarrhoea as a single primary outcome and find little evidence of an effect of the intervention, there is little we can infer about the intervention’s effectiveness as small relative effects with diarrhoea as an outcome can be compatible with larger reductions in enteric infection. Further to that, the intervention may have been very successful

as a behaviour change intervention. For example, caregivers might have adopted handwashing and water chlorination. But unbeknownst to the researchers the primary transmission pathway for enteric pathogens was geophagy or another alternative. The design of the methods of education was not at fault, it was the subject of the training that was poorly aligned with the context. The lack of contextualising information and observations from the causal chain between intervention and clinical outcome means there is little opportunity to triangulate evidence and interpret findings.

An update to the influential framework for designing and evaluating complex interventions by the UK's Medical Research Council was recently published.⁴¹ They identify four different but overlapping research perspectives and questions for complex interventions: efficacy, effectiveness, theory-based and systems' perspective. The latter three are of most relevance here, which we can summarise as: does the intervention produce the intended effects in real-world settings? What works in which circumstances and how? And, how do the system and intervention adapt to one another? We would argue that the 'intended effects' are often at several points in a causal pathway, such as changing behaviour, to reduce water and food contamination and, hence, the transmission of enteric pathogens and symptomatic illness (figure 2). It is, therefore, only by looking at these different outcomes that we can answer the effectiveness question, and in so doing start to answer the theory-based and systems questions. As the guidance describes, no trial provides a simple yes/no answer to the question 'did it work?', especially when the trial is examining interventions in complex systems.⁴¹

Trials of community interventions to tackle other infectious diseases can also provide useful exemplars for the WASH community. For example, a recent trial of mask wearing in 600 clusters incorporating over 340 000 people in Bangladesh used symptom reporting alongside seroprevalence studies and adherence measures, particularly mask wearing, as outcome measures.⁴² The researchers could both demonstrate an increase in mask wearing and a subsequent decrease in seroprevalence.

Most methods to correct for misclassification bias in an outcome require independent knowledge of the sensitivity and specificity of the measurement method. The misclassification errors associated with diarrhoea as an outcome are variable and difficult to estimate as we have described. While there are now methods that will handle misclassification error to produced unbiased estimates of treatment effects, it is always better still to improve the measurement.⁴³ Evidently, there may be a bias-variance trade-off to make: between a small, relatively uncertain but unbiased trial, or a large, 'certain' but biased one. However, we believe that the stronger consequence of our argument is that the reliance of a single outcome, particularly if it is diarrhoea, should be abandoned in favour of approaches that respect the complex nature of the intervention and system, and that allow for triangulation of the evidence across the causal pathway.

CONCLUSIONS

One explanation for the slow progress on reducing the risk of diarrhoeal disease in many LMICs may be that a solution, involving large-scale water and sewerage infrastructure, is unobtainable in many low-resource settings without significant external investment. In recent years, there have been many innovative technological solutions proposed for aspects of WASH like faecal sludge management, and access to clean drinking water and food preparation. Altogether, a package of such measures might provide significant relief in some settings that lack large-scale public health infrastructure. To identify what to include in such a successful programme, the question 'what works in WASH?' needs to be better answered using the best possible methods and measurements. We have argued that future trials in this area should not use survey-based diarrhoea as the primary outcome to avoid bias and inappropriate conclusions about the effects of an intervention. For a WASH intervention to be successful, it must cause a 'domino effect' across multiple mediating outcomes, such as behaviour change and interruption of pathogen transmission. Failure to reduce diarrhoea may or may not result from any one of these effects. Even if a trial were to demonstrate that an intervention causes a reduction in adverse clinical outcomes, even unbiased ones, the nature of the complex system means we may further struggle to generalise these findings and to other settings. Trial outcomes should be chosen from the causal pathway to better understand how an intervention functions or fails to do so, and in what context.

Twitter Samuel I Watson @siwatson

Contributors SIW and RL conceived the idea for the manuscript. SIW prepared the first draft. RL, TH and RTTR contributed to rewrites and edits.

Funding Medical Research Council (MR/V038591) and National Institute for Health Research (ARC West Midlands, EP/V028936).

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement There are no data in this work.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Richard J Lilford <http://orcid.org/0000-0002-0634-984X>

REFERENCES

- 1 Kotloff KL, Nataro JP, Blackwelder WC, *et al*. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the global enteric multicenter study, GEMs): a prospective, case-control study. *Lancet* 2013;382:209–22.
- 2 GBD 2019 Under-5 Mortality Collaborators. Global, regional, and national progress towards sustainable development goal 3.2 for neonatal and child health: all-cause and cause-specific mortality

- findings from the global burden of disease study 2019. *Lancet* 2021;398:870–905.
- 3 Wolf J, Hunter PR, Freeman MC, *et al*. Impact of drinking water, sanitation and handwashing with soap on childhood diarrhoeal disease: updated meta-analysis and meta-regression. *Trop Med Int Health* 2018;23:508–25.
 - 4 Cumming O, Arnold BF, Ban R, *et al*. The implications of three major new trials for the effect of water, sanitation and hygiene on childhood diarrhoea and stunting: a consensus statement. *BMC Med* 2019;17:173.
 - 5 Humphrey J, Pickering A, Null C, *et al*. The WASH benefits and shine trials. interpretation of findings on linear growth and diarrhoea and implications for policy: perspective of the investigative teams (P10-136-19). *Curr Dev Nutr* 2019;3.
 - 6 Luby SP, Rahman M, Arnold BF, *et al*. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. *Lancet Global Health* 2018;6:e302–15.
 - 7 Null C, Stewart CP, Pickering AJ, *et al*. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. *Lancet Glob Health* 2018;6:e316–29.
 - 8 Humphrey JH, Mbuya MNN, Ntozini R, *et al*. 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 Health* 2019;7:e132–47.
 - 9 World Health Organization. Diarrhoeal disease, 2022. Available: <https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease> [Accessed 8 Mar 2022].
 - 10 Rego RTT, Watson S, Gill P, *et al*. The impact of diarrhoea measurement methods for Under-Fives in low and middle income countries on reported diarrhoea rates: a systematic review and meta-analysis of methodological and primary empirical studies. *SSRN J* 2021.
 - 11 The DHS Program. Demographic and health surveys (DHS), 2018. Available: <https://dhsprogram.com/>
 - 12 Rego R, Watson S, Alam MAU, *et al*. A comparison of traditional diarrhoea measurement methods with microbiological and biochemical indicators: a cross-sectional observational study in the COX's Bazar displaced persons cAMP. *EClinicalMedicine* 2021;42:101205.
 - 13 Grenov B, Namusoke H, Nabukeera-Barungi N, *et al*. Validation of a simple stool diary used by caregivers to document diarrhea among young children in a low-income country. *J Pediatr Gastroenterol Nutr* 2017;65:156–64.
 - 14 Voskuijl W, Potani I, Bandsma R, *et al*. Stool frequency recording in severe acute malnutrition ('StoolSAM'); an agreement study comparing maternal recall versus direct observation using diapers. *BMC Pediatr* 2017;17:140.
 - 15 Schmidt W-P, Arnold BF, Boisson S, *et al*. Epidemiological methods in diarrhoea studies—an update. *Int J Epidemiol* 2011;40:1678–92.
 - 16 Whitworth HS, Gallagher KE, Howard N, *et al*. Efficacy and immunogenicity of a single dose of human papillomavirus vaccine compared to NO vaccination or standard three and two-dose vaccination regimens: a systematic review of evidence from clinical trials. *Vaccine* 2020;38:1302–14.
 - 17 RECOVERY Collaborative Group. Tocilizumab in patients admitted to hospital with COVID-19 (recovery): a randomised, controlled, open-label, platform trial. *Lancet* 2021;397:1637–45.
 - 18 Greenland S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol* 1996;25:1107–16.
 - 19 Neuhaus J. Bias and efficiency loss due to misclassified responses in binary regression. *Biometrika* 1999;86:843–55.
 - 20 Baker JM, Hasso-Agopowicz M, Pitzer VE, *et al*. Association of enteropathogen detection with diarrhoea by age and high versus low child mortality settings: a systematic review and meta-analysis. *Lancet Glob Health* 2021;9:e1402–10.
 - 21 Alam N, Henry FJ, Rahaman MM. Reporting errors in one-week diarrhoea recall surveys: experience from a prospective study in rural Bangladesh. *Int J Epidemiol* 1989;18:697–700.
 - 22 Boyer GS, Templin DW, Goring WP, *et al*. Discrepancies between patient recall and the medical record. Potential Impact on Diagnosis and Clinical Assessment of Chronic Disease. *Arch Intern Med* 1995;155:1868–72.
 - 23 Rego R, Watson S, Ishengoma P, *et al*. Effectiveness of SMS messaging for diarrhoea measurement: a factorial cross-over randomised controlled trial. *BMC Med Res Methodol* 2020;20:174.
 - 24 Eisenberg JNS, Scott JC, Porco T. Integrating disease control strategies: balancing water sanitation and hygiene interventions to reduce diarrheal disease burden. *Am J Public Health* 2007;97:846–52.
 - 25 Eisenberg JNS, Trostle J, Sorensen RJD, *et al*. Toward a systems approach to enteric pathogen transmission: from individual independence to community interdependence. *Annu Rev Public Health* 2012;33:239–57.
 - 26 Mullineaux-Sanders C, Suez J, Elinav E, *et al*. Sieving through gut models of colonization resistance. *Nat Microbiol* 2018;3:132–40.
 - 27 Pickard JM, Núñez G. Pathogen colonization resistance in the gut and its manipulation for improved health. *Am J Pathol* 2019;189:1300–10.
 - 28 Riggs MM, Sethi AK, Zabarsky TF, *et al*. Asymptomatic carriers are a potential source for transmission of epidemic and nonepidemic *Clostridium difficile* strains among long-term care facility residents. *Clin Infect Dis* 2007;45:992–8.
 - 29 Kariuki S, Dyson ZA, Mbae C, *et al*. Multiple introductions of multidrug-resistant typhoid associated with acute infection and asymptomatic carriage, Kenya. *Elife* 2021;10:67852. doi:10.7554/eLife.67852
 - 30 Pearl J. The causal mediation formula—a guide to the assessment of pathways and mechanisms. *Prev Sci* 2012;13:426–36.
 - 31 Aragie S, Tadesse W, Dagneu A, *et al*. Changing hygiene behaviours: a cluster-randomized trial, Ethiopia. *Bull World Health Organ* 2021;99:762–72.
 - 32 Freeman MC, Ellis AS, Ogutu EA, *et al*. Impact of a demand-side integrated WASH and nutrition community-based care group intervention on behavioural change: a randomised controlled trial in Western Kenya. *BMJ Glob Health* 2020;5:e002806.
 - 33 Lilford RJ, Chilton PJ, Hemming K, *et al*. Evaluating policy and service interventions: framework to guide selection and interpretation of study end points. *BMJ* 2010;341:c4413.
 - 34 Vaz Nery S, Traub RJ, McCarthy JS, *et al*. WASH for WORMS: a cluster-randomized controlled trial of the impact of a community integrated water, sanitation, and hygiene and deworming intervention on soil-transmitted helminth infections. *Am J Trop Med Hyg* 2019;100:750–61.
 - 35 Knee J, Sumner T, Adriano Z, *et al*. Effects of an urban sanitation intervention on childhood enteric infection and diarrhea in Maputo, Mozambique: a controlled before-and-after trial. *Elife* 2021;10:62278. doi:10.7554/eLife.62278
 - 36 Rogawski McQuade ET, Platts-Mills JA, Gratz J, *et al*. Impact of water quality, sanitation, handwashing, and nutritional interventions on enteric infections in rural Zimbabwe: the sanitation hygiene infant nutrition efficacy (SHINE) trial. *J Infect Dis* 2020;221:1379–86.
 - 37 Capone D, Berendes D, Cumming O, *et al*. Analysis of fecal sludges reveals common enteric pathogens in urban Maputo, Mozambique. *Environ Sci Technol Lett* 2020;7:889–95.
 - 38 Diggle PJ. Estimating prevalence using an imperfect test. *Epidemiol Res Int* 2011;2011:1–5.
 - 39 Luby SP, Rahman M, Arnold BF, *et al*. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. *Lancet Glob Health* 2018;6:e302–15.
 - 40 McShane BB, Gal D, Gelman A, *et al*. Abandon statistical significance. *Am Stat* 2019;73:235–45.
 - 41 Skivington K, Matthews L, Simpson SA, *et al*. A new framework for developing and evaluating complex interventions: update of medical research council guidance. *BMJ* 2021;374:n2061.
 - 42 Abaluck J, Kwong L, Styczynski A, *et al*. The impact of community masking on COVID-19: a cluster-randomized trial in Bangladesh, 2021. Available: <https://www.poverty-action.org/publication/impact-community-masking-covid-19-cluster-randomized-trial-bangladesh>
 - 43 Liu H, Zhang Z. Logistic regression with misclassification in binary outcome variables: a method and software. *Behaviormetrika* 2017;44:447–76.