The Impact of Diarrhoea Measurement Methods for Under-Fives in Low and Middle Income Countries on Estimated Diarrhoea Incidence Rates at the Population Level: *A Systematic Review and Meta-Analysis of Methodological and Primary Empirical Studies*

Ryan Rego[1], Samuel Watson[2], Paramjit Gill[1], Richard Lilford[2]

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

2 Institute for Applied Health Research, University of Birmingham, Birmingham, United Kingdom

Sustainable Development Goal: Good Health and Wellbeing

Abstract

Objective: We systematically reviewed all studies published between 2000 and June 2021 that estimated under-five diarrhoea rates in low and middle income countries and extracted data on diarrhoea rates, measurement methods, and reactivity.

Methods: We summarised data from studies that performed direct comparisons of methods, and indirectly compared studies which utilised only one method using meta-regression to determine the association between methods and estimated diarrhoea rates.

Results: 296 studies met our inclusion criteria: four direct comparisons and 288 studies utilising only one measurement method. Meta-regression across all studies showed that diarrhoea rates were sensitive to method of measurement. We estimated that passive surveillance methods were associated with a 97% lower estimated rate than active surveillance(IRR=0·03,95%CI[0·02,0·06]). Among active surveillance studies, a doubling of recall period was associated with a 48% lower rate(IRR=0·52[0·46,0·60]), while decreased questioning frequency was associated with a higher estimated rate: at the extreme, one time questioning yielded an over 4X higher rate than daily questioning(IRR=4·22[2·73,6·52]).

Conclusions: Estimated diarrhoea rates are sensitive to their measurement methods. There is a need for a standardisation of diarrhoea measurement methods, and for the use of other outcomes in the measurement of population level gastrointestinal health.

Keywords: Diarrhoea; Surveillance; Epidemiology; Child Health; WASH

Introduction

Effective surveillance of diarrhoea in children under-five at the population level is required to track outbreaks, allocate public health resources, and evaluate water, sanitation, and hygiene (WASH) interventions(1). However, the choice of method used to ascertain whether an episode of diarrhoea has occurred has been hypothesized to impact estimated diarrhoea rates(2, 3). If true, the results of observational and intervention studies may differ by method of measurement, thereby obscuring the effects of the explanatory variable(s) of interest. Since diarrhoea is one of the

biggest causes of death in children, this is a methodological point of considerable practical importance, for surveillance, evaluation of interventions, and establishing the true burden of diarrhoea associated morbidity and mortality. The two most common methods of diarrhoea surveillance at the population level, passive and active surveillance, take different approaches. Passive surveillance relies on data collected from health facilities and therefore excludes all children with diarrhoea who do not attend a facility. Passive surveillance estimates are therefore skewed towards more severe disease and away from marginalised groups such as slum-dwellers, refugees, and migrants, who are less likely to visit health facilities and who are more likely to visit informal facilities than non-marginalised groups(4-6) (7-9). Passive surveillance is a useful, inexpensive tool to detect new outbreaks of severe diseases such as cholera, but is arguably less useful as an epidemiological tool for the measurement of population-level diarrhoea rates or in trials of WASH interventions.

Active surveillance, based on door-to-door surveys, provides a more complete report of diarrhoea rates than passive surveillance, but may also be subject to measurement error and bias(4). Carers may forget events that happened in the past, particularly during longer lengths of recall(10, 11). They may also have a poor understanding of what diarrhoea is(12). UNICEF and the Demographics and Health Surveys (DHS) program have recommended a method based on asking carers if their child has had three or more loose or watery stools in any 24-hour period within the previous 14 days(13). However, this method is by no means universally applied. In addition to concerns over measurement error, concern has been expressed that diarrhoea rates may be subjective to bias due to 'reactivity,' where despite any true clinical difference, people report different diarrhoea rates for psychological reasons. For example, they may be the beneficiaries of an intervention and not want to appear ungrateful; not want to answer in a way that is not socially desirable; or be guided subliminally or non-verbally by the surveyors (e.g. if the question is asked in a way which steers the respondents towards a certain answer) (14, 15). As such, reactivity creates a particular concern for evaluations of WASH interventions.

In order to examine the association between the method used in diarrhoea measurement and diarrhoea rates we conducted a systematic review of all studies published between 2000 and June 2021 that report under-five diarrhoea rates in low and middle income countries (LMICs). We examined studies that perform direct 'head-to-head' comparisons of different methods in order to estimate differences in diarrhoea rates by method. However, we found only four such studies. We therefore obtained studies that used only one measurement method so that we could compare the estimated diarrhoea rates of each method across studies indirectly by means of meta-analytical methods(16, 17).

The aims of the systematic review and meta-analysis were to determine: (1) the frequency of use of the different diarrhoea measurement methods; (2) the association between passive and active surveillance methods and estimated diarrhoea rates; (3) the association between recall periods, questioning frequencies, and prospective (diary) vs retrospective recall on estimated diarrhoea rates among active surveillance studies; and (4) the extent of reactivity in diarrhoea measurement.

Methods

Search Strategy and Selection Criteria

We conducted a systematic review of studies published between 2000 and June 2021 that made quantitative measurements of diarrhoea rates among children under the age of five in LMICs (as defined by OECD)(18). The search strategy aimed to capture any study that estimated diarrhoea rates among under-fives, including both studies that

performed direct ' head-to-head' comparisons of methods, and studies estimating diarrhoea rates using only one method that we can compare indirectly. As we are only interested in population level diarrhoea rates, we excluded studies that were not designed to capture population level diarrhoea rates, such as studies which measured hospital acquired infection; clinical trials in which diarrhoea was an adverse drug event; and case-control studies in which diarrhoea was the case. Importantly, these exclusions do not include WASH trials which took place in the community setting. Studies were restricted to English or French (Table 1).

Table 1: Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Reports diarrhoea rates among under-fives (i.e. incidence or prevalence)	Measures Hospital-Acquired Infection
Takes place in LMICs	Clinical trials in which diarrhoea was an adverse drug event
Published between 2000 and June 2021	Case-control studies in which diarrhoea is the case
English or French language	

We searched the MEDLINE, Embase, and PubMed databases for studies matching the inclusion and exclusion criteria. The search string (Appendix 1) included key words relating to diarrhoea and population-level disease measurement. The string then restricted the results to human studies and studies in LMICs.

SW and RR independently screened each title and abstract, and any disagreements were resolved through discussion with RL. Full texts were screened in a similar manner. Where full texts were not available, we requested the article from the University of Warwick's Article Reach Service. We excluded unavailable studies and duplicate studies. In the event that multiple studies used the same data source, we selected a random study for inclusion.

Data Extraction

RR extracted the data, with SW duplicating data extraction for random 15% of the included studies. The random 15% extracted by SW matched completely with the extraction by RR. As a single reported study may have included several independent reports of diarrhoea rates, we treated each report as separate 'observations' within one study. This would apply to all the direct studies but also arose in indirect studies when conducted in more than one site, included multiple rounds of data collection and/or was a trial with multiple arms. For example, a two-armed trial which estimated diarrhoea rates at base-line and end-line yielded four observations.

Data extracted (Table 2) included participant demographics, study design, if the study was an observational study using primary data (including non-randomised trials), an observational study using secondary data or a randomised control trial; diarrhoea rates; measurement methods; and if the study was a direct comparison of methods. We defined direct comparison studies as those which included at least two separate arms, each with a different method of diarrhoea measurement (including altering recall period or questioning frequency) that compare estimated diarrhoea rates between each arm.

Table 2: Data extracted for each observation

Variable	Units (if applicable)	Categorisation (if categorised)
Authors	NA	NA
Study Title	NA	NA
Year of Publication	Years	NA
Year of Data Collection	Years	NA
Mean Participant Age	Years (with Standard Error)	NA
Participant Sex Breakdown	% Female	NA
Geography	Urban, Rural, or Mixed	NA
Country	NA	Regions as defined by the United Nations Development Program (UNDP)(19)
Intervention (for randomised control trials (RCTs))	NA	WASH, Health, or Nutrition
Effectiveness of Intervention (for RCTs)	Incidence Risk Ratio	NA
Diarrhoea Rate (in incidence)	Episodes per child-year (converted if necessary)	NA
Measurement Type	NA	Passive, Retrospective Active, Prospective (Diary) Active
Recall Period	Days	NA
Questioning Frequency	Days	Daily, Weekly, Monthly, or Annually or longer
Questioning Type	NA	Pictorial, Verbal, or other
Sample Size (observed cases for passive surveillance studies)	NA	NA
Study Design	NA	Observational (Primary), Observational (Secondary), RCT (broken down by arm)
Direct Comparison of Methods	NA	Yes/No

Data Analysis

We indirectly compared the diarrhoea rates from included observations, including those from studies that performed direct comparisons of measurement methods, and studies using only a single method of measurement. We summarised the key variables in each individual observation, including estimated diarrhoea rates, measurement type, year, and region (as defined by UNDP). We estimated a hierarchical meta-regression model for log diarrhoea rate (incidence in episodes per child-year), adjusting for region and time (in years) and interactions between time and region. We also adjusted for study design, including a categorical variable with levels: (1) Observational studies which use primary data sources; (2) observational studies which use secondary data sources; (3) RCT intervention arms before the intervention, or the control arm (if reported); and (4) RCT experimental arm after the intervention. The model was estimated in StataSE Version 15 using generalised least squares with random effects at the study level, to account for within-study correlation, and weighting by the study sample size(19, 20).

We estimated two separate meta-regression models. The first iteration included the observations from both passive and active surveillance studies to estimate the effect of surveillance types (passive/active) on estimated

diarrhoea rates, and as such included a dummy variable for surveillance type. We also estimated pooled temporal and regional trends from this model.

The second iteration included only observations from active surveillance studies to examine the effects of variables exclusive to active surveillance studies on estimated diarrhoea rates. These included variables for recall period (as a continuous numeric term in days), and questioning frequency and recall type (both as categorical variables). Further, any other variations found between active surveillance studies, including reactivity and questioning type (e.g. verbal or pictorial) were included.

Results

Study Identification

We identified 2,040 studies in total, which was reduced to 1,973 after duplicates were removed. Abstract and title screening yielded 577 studies, with a further 289 excluded after full-text review. Common reasons for exclusion included not presenting data on under-fives (n=64), not including data on diarrhoea rates (n=58), and not being able to obtain the full text (n=4). Thus 288 full text studies (Appendix 2) were included in the final review (Figure 1).

Figure 1: PRISMA Diagram of the Systematic Review

Study Characteristics

We first describe the characteristics of the described studies, such as observations per study, data collection method, geography, and others. As stated, many studies included more than one observation of diarrhoea rate, arising from observations at different time points, being a trial with two or more arms, or data collection in more than one location. Appendix 3 presents information on the number of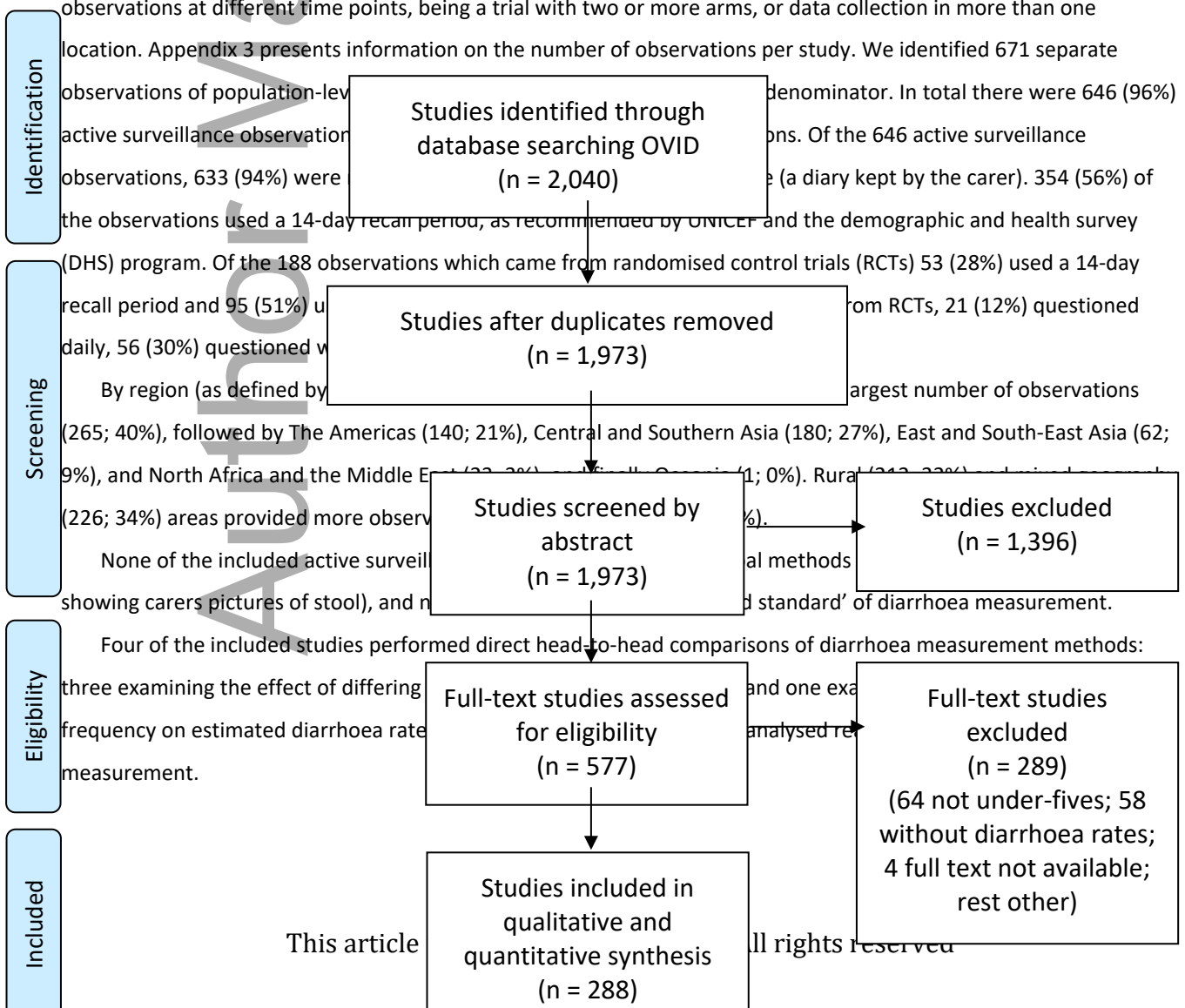 observations per study. We identified 671 separate observations of population-lev[...] denominator. In total there were 646 (96%) active surveillance observation[...]ns. Of the 646 active surveillance observations, 633 (94%) were [...] (a diary kept by the carer). 354 (56%) of the observations used a 14-day recall period, as recommended by UNICEF and the demographic and health survey (DHS) program. Of the 188 observations which came from randomised control trials (RCTs) 53 (28%) used a 14-day recall period and 95 (51%) u[...]rom RCTs, 21 (12%) questioned daily, 56 (30%) questioned w[...]

By region (as defined by [...] argest number of observations (265; 40%), followed by The Americas (140; 21%), Central and Southern Asia (180; 27%), East and South-East Asia (62; 9%), and North Africa and the Middle E[...] Oceania (1; 0%). Rura[...] (226; 34%) areas provided more observ[...]%).

None of the included active surveil[...]al methods showing carers pictures of stool), and n[...]d standard' of diarrhoea measurement.

Four of the included studies performed direct head-to-head comparisons of diarrhoea measurement methods: three examining the effect of differing [...] and one exa[...] frequency on estimated diarrhoea rate[...] analysed re[...] measurement.

Identification

Screening

Eligibility

Included

Studies identified through database searching OVID (n = 2,040)

Studies after duplicates removed (n = 1,973)

Studies screened by abstract (n = 1,973)

Studies excluded (n = 1,396)

Full-text studies assessed for eligibility (n = 577)

Full-text studies excluded (n = 289) (64 not under-fives; 58 without diarrhoea rates; 4 full text not available; rest other)

Studies included in qualitative and quantitative synthesis (n = 288)

Table 3: Summary table of the 671 observations included in the systematic review, with average estimated diarrhoea rates for key variables

		Passive n(% of total)	Prospective (Diary) Active n(% of total)	Retrospective Active n(% of total)	Total	Average estimated diarrhoea rate (Episodes per Child-Year)
	TOTAL	25	13	633	671	6·75
Data Collection Year	2016-2018	1(2.9%)	0(0%)	33(97.1%)	34	7.08
	2011-2015	12(6·8%)	2(1·1%)	163(92.1%)	177	6.63
	2006-2010	6(3·2%)	2(1·1%)	180(95·7%)	188	6.60
	2001-2005	4(3·4%)	2(1·7%)	112(94·9%)	118	7.80
	1996-2000	2(2·5%)	1(1·3%)	77(96·2%)	80	6.15
	1991-1995	0(0%)	0(0%)	38(100%)	38	5.99
	1986-1990	0(0%)	3(42·9%)	4(57·2%)	7	3.54
	Missing Data	0(0%)	3(10·3%)	26(89·7%)	29	7.46
Region	Sub-Saharan Africa	5(1·9%)	3(1·1%)	257(97.0%)	265	6.32
	North Africa and the Middle East	0(0%)	0(0%)	23(93·4%)	23	6.56
	Central and Southern Asia	3(1·7%)	2(1·1%)	175(97·2%)	180	9.41
	Eastern and South-East Asia	13(21·0%)	4(6·4%)	45(72·6%)	62	3.20
	Oceania	0(0%)	0(0%)	1(100·0%)	1	4.02
	The Americas	5(3·6%)	3(2·1%)	132(94·3%)	140	5.74
Study Design	Observational (Primary)	19(6·5%)	11(3·7%)	264(89·8%)	294	8·21
	Observational (Secondary)	6(3·2%)	0(0%)	183(96·8%)	189	3·87

	RCT (Baseline or Control)	0(0%)	1(1·1%)	93(98·9%)	94	8·37
	RCT (Experimental Post-Intervention)	0(0%)	1(1·1%)	93(98·9%)	94	6.32
Geography	Mixed	14(5·8%)	2(0·8%)	226(93·4%)	242	4.02
	Rural	4(1·8%)	4(1·8%)	212(96·3%)	220	10.55
	Urban	4(3·1%)	7(5·5%)	116(91·3%)	127	5.15
	Missing Data	3(3·7%)	0(0%)	79(96·4%)	82	6.99
Questioning Frequency	Daily	NA	0(0%)	55(100%)	55	4.68
	Weekly	NA	9(9·5%)	86(90·5%)	95	4.49
	Monthly	NA	1(1·3%)	77(98·8%)	78	13.47
	Annual+	NA	0(0%)	30(100%)	30	5·43
	One Off	NA	3(0·8%)	385(99·2%)	388	6.76
Recall Period	1-3 days	NA	1(1·6%)	60(98·4%)	61	25.88
	4-14 days	NA	11(2·1%)	520(97·9%)	531	5.22
	15-30 days	NA	1(2·6%)	37(97·4%)	38	3.94
	31-90 days	NA	0(0%)	13(100%)	13	1.00
	91+ days	NA	0(0%)	3(12·5%)	3	1.83

Table 4: A random effects model showing the association between study characteristics and estimated diarrhoea rates for all observations and only active surveillance observations .

		Active Vs Passive		Within Active	
		IRR	95%CI	IRR	95%CI
Surveillance Type	Reference: Active Surveillance	Ref.	Ref.	Ref.	Ref.
	Passive Surveillance	0·03	(0·02,0·06)	NA	NA
Region	Reference: Sub-Saharan Africa	Ref.	Ref.	Ref.	Ref.
	North Africa/Middle East	1·31	(0·44,3·87)	1.32	(0.45,3.86)
	Central and South Asia	0·88	(0·13,6·14)	0.87	(0.13,6)
	Eastern and South-East Asia	1·21	(0·47,3·07)	1.2	(0.47,3.07)

	Oceania	1·12	(1·05,1·19)	1.11	(1.05,1.17)
	South America and Caribbean	0·63	(0·47,0·84)	0.67	(0.51,0.9)
Year	Year	0·97	(0·96,0·98)	0.97	(0.97,0.98)
Region and Year	Reference: Sub-Saharan Africa	Ref.	Ref.	Ref.	Ref.
	North Africa/Middle East	0.99	(0.95,1·02)	0.99	(0.95,1.02)
	Central and South Asia	0.99	(0.93,1·06)	0.99	(0.93,1.06)
	Eastern and South-East Asia	0.99	(0.96,1·02)	0.99	(0.96,1.02)
	Oceania	1.00	(1.00,1.00)	1.00	(1.00,1.00)
	The Americas	1·02	(1·01,1·03)	1.02	(1.01,1.03)
Study Design	Reference: Observational Primary	Ref.	Ref.	Ref.	Ref.
	Observational Secondary	1·05	(0·89,1·22)	1.04	(0.89,1.22)
	RCT (Baseline or Control)	1·19	(0·85,1·67)	1.5	(1.12,2.02)
	RCT (Experimental Post-Intervention)	1·04	(0·75,1·43)	1.33	(0.98,1.79)
	Log of Measurement Period	NA	NA	0·52	(0·46,0·60)
Questioning Frequency	Reference: Daily	NA	NA	NA	NA
	Weekly	NA	NA	1.24	(0.77,1.99)
	Monthly	NA	NA	2.57	(1.36,4.84)
	Annually or Longer	NA	NA	4.8	(2.73,8.45)
	One Off	NA	NA	4.22	(2.73,6.52)
Method	Reference: Diary	NA	NA	Ref.	Ref.
	Self Report	NA	NA	0.85	(0.53,1.38)
	Constant	7·98	(5·72,11·14)	13.86	7.3
Random Effects Parameters	Variance (Constant)	1·07	(0·86,1·33)	0·67	(0·47,0·95)
	Variance (Residual)	0·08	(0·06,0·11)	0·08	(0·05,0·11)

Differences between Active and Passive Surveillance

As stated in the introduction, the two main categories of disease surveillance are active surveillance (community surveying) and passive surveillance (measures of visits to health facilities). No studies performed direct 'head-to-head' comparisons of active and passive surveillance. After model-based adjustment to perform an indirect comparison of passive and active surveillance, passive surveillance was associated with a 97% lower estimated diarrhoea rate than active surveillance (Incidence Risk Ratio (IRR)=0·03,95%CI[0·02,0·06]) (Table 4).

Impacts of Factors Within Active Surveillance Studies on Estimated Diarrhoea Rates

Specific to active surveillance studies, questioning frequency (how often participants are surveyed), recall period of diarrhoea surveying, and retrospective vs prospective questioning may differ.

Questioning Frequency

One study directly examined the effect of differing questioning frequencies on estimated diarrhoea rates; Zwane _et al_ (2011) estimated that biweekly surveys had a 7-15% lower diarrhoea rate than six-monthly surveys when using the same recall period(21). Our indirect comparison of active surveillance observations produced comparable results; after model-based adjustment, we found that that less frequent questioning was associated with an increase in estimated diarrhoea rates. For example, one-time questioning was associated with a rate over four times higher than daily questioning (IRR=4·22[2·73,6·52]) (Table 4). This, however, is not evident graphically in unadjusted crude data – however, a large amount of variance as questioning frequency increases can still be seen (Figure 2a).

Recall Period

Three studies directly examined the effect of differing recall periods on estimated diarrhoea rates, although the recall periods examined were different. Melo _et al_ (2007) found that diarrhoea rates were cut by a third when carers recall over 4 weeks vs. 24 hours(22). Feikin _et al_ (2010) similarly estimated that diarrhoea rates were cut by a fifth when carers recall over 11-13 days vs. 1-2 days(10). Lee _et al_ (2010), estimated that estimated diarrhoea rates were similar for carers who recalled over a 72-hour period vs. a 24-hour period(23), but this is a much shorter range than that investigated in the other two studies.

Based on an indirect comparison of the included active surveillance observations, we estimated that recall periods and estimated diarrhoea rates were inversely associated. After model-based adjustment, we found that a doubling of recall period was associated with a 48% reduction in diarrhoea rate (IRR=0·52[0·46,0·60]; Table 4). This is also evident graphically in the crude data (Figure 2b).

Prospective vs Retrospective

No studies directly compared prospective (diary) recall designs against retrospective. We estimated through indirect comparison that retrospective recall observations were associated with a lower rate than prospective (diary) observations, but the effect size was relatively uncertain (IRR=0·85[0·53,1·38]; Table 4).

Figure 2a: Box plot of estimated diarrhoea rate against questioning frequency for active surveillance studies
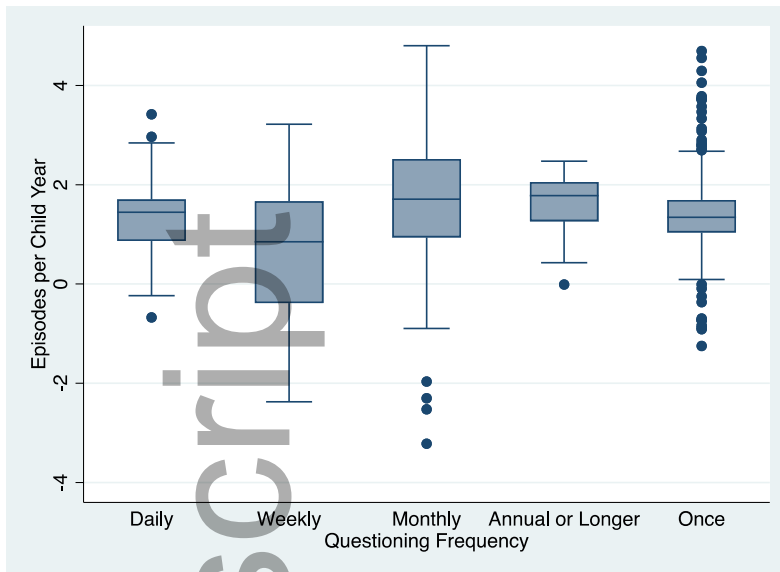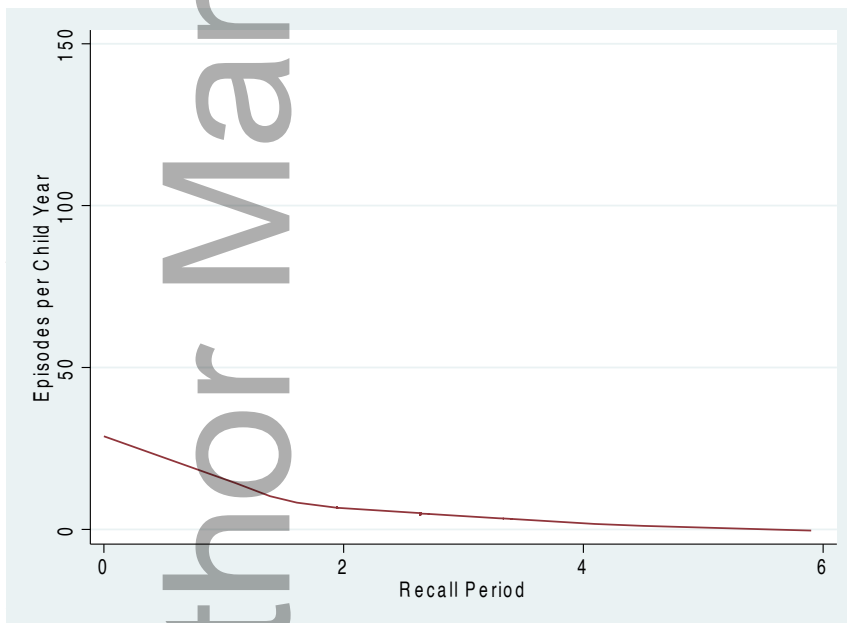
Figure 2b: Scatter plot and trendline of estimated diarrhoea rates against recall period for active surveillance studies

... the methods used in their

... tive surveillance, as well as factors

... tes than active surveillance methods.

... ajority of cases where, if asked, they

... ess to healthcare among infants in

... n diarrhoea is influenced by diarrhoea

... (5-9).

... g frequencies influence estimated

... gher questioning frequencies: one-off

... a rate than daily questioning. We also

... diarrhoea rates: a doubling of recall

period was associated with a halving of estimated diarrhoea rate.

Factors that Result in Subjective Measurements during Active Surveillance

As the distinction between and recall of a diarrhoeal or non-diarrhoeal stool by carers is largely subjective, several cognitive factors can affect measurement. These include respondent fatigue (becoming tired of answering questions), recall bias (forgetting events that have occurred in the past), perception bias (not understanding the question being asked), and reactivity (answering differently due to experiencing an intervention).

Respondent Fatigue

Declining diarrhoea rates with increasing questioning frequency (but the same recall period) suggest respondent

fatigue – participants may be inclined to pay attention to their bowel movements at first but lose motivation with further rounds of questioning.

Recall Bias

Recall bias is the effect of forgetting: participants are more likely to recall recent than older events. We would expect a lower reported number of diarrhoea episodes with longer recall periods and this was borne out by our analysis including two of the three head-to-head comparisons – the exception examined a much smaller gap between questions than the other two(10, 24) This finding was corroborated by a more recent head-to-head comparison where daily recall was associated with a 30 percentage point higher estimated diarrhoea rate than fortnightly recall during a text message survey of under-five diarrhoea in urban Tanzania(25). While not examined in our review, it has also been reported that the effect of recall bias is more apparent for moderate diarrhoea compared to severe diarrhoea. Zafar *et al (*2010), for example, found that moderate diarrhoea is reported at half the rate of severe diarrhoea during longer recall periods(11).

Other factors affecting diarrhoea measurement

Other factors outside the scope of this review can further influence estimated diarrhoea rates. For example, poor caregiver perception of diarrhoea (understanding what is or is not diarrhoea) can result in error in diarrhoea measurement. Voskuijl *et al* (2017) determined that carers of children under five were only able to identify 56%-75% of loose or watery stool, and 80% healthy stools(12).

   Another relevant phenomenon is "reactivity" whereby participants adjust their answers to a survey according to how they believe they ought to respond, regardless of any true underlying difference. We did not identify any studies of reactivity in this review, but it has been discussed as a potential explanatory factor in previous trials. Luby *et al* (2018) mentioned courtesy bias as a potential source of reactivity, stating 'people who received the intervention might have been grateful and, out of courtesy, reported less diarrhoea' (14). Wood *et al* (2008) further found evidence for reactivity in clinical trials for various diseases, reporting that inadequate concealment of interventions is associated with improved treatment performance in trials, particularly for subjective outcomes(26). It was not possible to examine reactivity through courtesy bias or inadequate concealment in our review as WASH trials by nature are unblinded.

Conclusion

The magnitude of the variation in diarrhoea rates, even among active methods of surveillance, suggests the need for standardization of diarrhoea measurement methods to facilitate comparisons between studies. Despite the fourteen day UNICEF and DHS standard, there does seem to be a trend towards using a 7-day recall period. It was the most frequently used recall period (51%) among RCTs in our review. Three of the three recent large integrated WASH trials (the SHINE trials in Bangladesh and Kenya, and the WASH Benefits trial in Zimbabwe) also used a seven day retrospective recall to measure diarrhoea(27-30), in contravention of the UNICEF and DHS guidelines. However, the above three trials differed among themselves with respect to question frequency: the SHINE trials questioned carers annually, while the WASH Benefits trial questioned mothers every 'two to six' months(28-30)

   It could be argued that lack of standardisation simply introduces measurement error in trials that can be counteracted by increasing sample size. However, this is only likely to be true if it is assumed that the different

methods affect only the propensity of someone with a true case of diarrhoea (or indeed enteric infection) to report a case of diarrhoea (the "sensitivity" of the method)(31). If, however, there is a loss of "specificity" – the propensity of someone who did not have diarrhoea (or enteric infection) to report a case of diarrhoea – then intervention effects will also be biased across studies using different methods. It is also likely that any measurement errors will bias results towards the null, rather than towards reports of intervention effectiveness(31). It is therefore possible that the choice of methodology is at least partially responsible for the widely varying, and often disappointing, results of evaluations on WASH interventions(14, 28, 30, 32, 33).

While a widely accepted standard would facilitate comparisons across different observational and experimental studies, this raises the question of what the optimal standard may be that would also produce reliable reports of diarrhoea rates. There is no "gold standard" method for measurement of diarrhoea rates. In part, this is because of the difficulty in defining the underlying construct and providing a culturally and linguistically consistent definition of a case or episode of 'diarrhoea'. Direct observation by an expert might constitute a gold standard against which other methods could be compared, as has been described above in the study by Voskuiljl and colleagues (2017)(12). However, judgements among experts may not be universal. Moreover, the collection of every stool and use of experts to classify them quickly becomes impractical at larger scales.

We propose two policies to mitigate the problem. First, an agreed consensus method for measurement of diarrhoea rates in surveys. Second, triangulation of diarrhoea rates with other observations that reflect on gastrointestinal health when interventions are evaluated. Many WASH Studies already include anthropometric measurements as outcomes alongside diarrhoea. Further, direct measurement of environmental contamination and pathogen levels in stool samples should complement diarrhoea rates in clinical studies. This would also allow for determination of how much diarrhoea is attributable to infection (and which can be reduced by WASH interventions), rather than non-infective reasons (which would likely not be impacted by WASH interventions)(33). Investigation of the link between interventions, environmental contamination, and the profiles of pathogen carriage in childhood stools is an important topic for scientific research.

Funding

References

1.   (WHO) WHO. Diarrhoea Disease Geneva: World Health Organization; 2017 [Available from: https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease.

2.   Watson SI, Sartori J, Uthman O, Lilford RJ. Health effects of sanitation facilities: a Bayesian semiparametric analysis of compositional data. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2019.

3.   Wolf J, Hunter PR, Freeman MC, Cumming O, Clasen T, Bartram J, et al. Impact of drinking water, sanitation and handwashing with soap on childhood diarrhoeal disease: updated meta-analysis and meta-regression. Tropical medicine & international health. 2018;23(5):508-25.

4.   McNabb SJ, Conde JM, Ferland L, Macwright W, Memish Z, Okutani S, et al. Transforming Public Health

Surveillance-E-Book: Proactive Measures for Prevention, Detection, and Response: Elsevier; 2016.

5. Troeger C, Blacker BF, Khalil IA, Rao PC, Cao S, Zimsen SR, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Infectious Diseases. 2018;18(11):1211-28.

6. Thiam S, Diene AN, Sy I, Winkler MS, Schindler C, Ndione JA, et al. Association between Childhood Diarrhoeal Incidence and Climatic Factors in Urban and Rural Settings in the Health District of Mbour, Senegal. International journal of environmental research and public health. 2017;14(9).

7. Azage M, Kumie A, Worku A, C Bagtzoglou A, Anagnostou E. Effect of climatic variability on childhood diarrhea and its high risk periods in northwestern parts of Ethiopia. PloS one. 2017;12(10):e0186933.

8. Norredam M. Migrants' access to healthcare. Danish medical bulletin. 2011;58(10):B4339-B.

9. Hjortsberg C. Why do the sick not utilise health care? The case of Zambia. Health economics. 2003;12(9):755-70.

10. Feikin DR, Audi A, Olack B, Bigogo GM, Polyak C, Burke H, et al. Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. International journal of epidemiology. 2010;39(2):450-8.

11. Zafar SN, Luby SP, Mendoza C. Recall errors in a weekly survey of diarrhoea in Guatemala: determining the optimal length of recall. Epidemiology and infection. 2010;138(2):264-9.

12. Voskuijl W, Potani I, B, sma R, Baan A, White S, et al. Stool frequency recording in severe acute malnutrition ('StoolSAM'); an agreement study comparing maternal recall versus direct observation using diapers. 2017;17(1):140.

13. Larsen DA, Grisham T, Slawsky E, Narine L. An individual-level meta-analysis assessing the impact of community-level sanitation access on child stunting, anemia, and diarrhea: Evidence from DHS and MICS surveys. PLOS Neglected Tropical Diseases. 2017;11(6):e0005591.

14. Luby SP, Rahman M, Arnold BF, Unicomb L, Ashraf S, Winch PJ, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. 2018;6(3):e302-e15.

15. Gizaw Z, Woldu W, Bitew BD. Child feeding practices and diarrheal disease among children less than two years of age of the nomadic people in Hadaleala District, Afar Region, Northeast Ethiopia. International breastfeeding journal. 2017;12(1):24.

16. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339:b2535.

17. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of Observational Studies in EpidemiologyA Proposal for Reporting. JAMA. 2000;283(15):2008-12.

18. Development OfEC-oa. DAC List of ODA Recipients: Organisation for Economic Co-operation and Development; 2018 [Available from: https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/DAC_List_ODA_Recipients2018to2020_flows_En.pdf.

19. Division UNS. SDG Indicators Regional groupings used in 2017 Report and Statistical Annex New York: United Nations; 2017 [Available from: https://unstats.un.org/sdgs/indicators/regional-groups/.

20. StataCorp. StataSE Version 15. StataCorp; 2017.

21. Zwane AP, Zinman J, Van Dusen E, Pariente W, Null C, Miguel E, et al. Being surveyed can change later behavior and related parameter estimates. Proceedings of the National Academy of Sciences. 2011;108(5):1821-6.

22. Melo MCNd, Taddei JAdAC, Diniz-Santos DR, May DS, Carneiro NB, Silva LR. Incidence of diarrhea: poor parental recall ability. The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases. 2007;11(6):571-9.

23. Lee G, Cama V, Gilman RH, Cabrera L, Saito M, Checkley W, et al. Comparison of two types of epidemiological surveys aimed at collecting daily clinical symptoms in community-based longitudinal studies. 2010;20(2):151-8.

24. Melo MCNd, Taddei JAdAC, Diniz-Santos DR, May DS, Carneiro NB, Silva LR. Incidence of diarrhea: poor parental recall ability. 2007;11(6):571-9.

25. Rego R, Watson S, Ishengoma P, Langat P, Otieno HP, Lilford R. Effectiveness of SMS messaging for diarrhoea measurement: a factorial cross-over randomised controlled trial. BMC Medical Research Methodology. 2020;20(1):174.

26. Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ. 2008;336(7644):601.

27. Cumming O, Curtis V. Implications of WASH Benefits trials for water and sanitation. Lancet Glob Health. 2018;6(6):e613-e4.

28. Humphrey JH, Mbuya MNN, Ntozini R, Moulton LH, Stoltzfus RJ, Tavengwa NV, 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(1):e132-e47.

29. Luby SP, Rahman M, Arnold BF, Unicomb L, Ashraf S, Winch PJ, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. The Lancet Global Health. 2018;6(3):e302-e15.

30. Null C, Stewart CP, Pickering AJ, Dentz HN, Arnold BF, Arnold CD, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. The Lancet Global Health. 2018;6(3):e316-e29.

31. GREENLAND S. Basic Methods for Sensitivity Analysis of Biases. International Journal of Epidemiology. 1996;25(6):1107-16.

32. Pickering AJ, Null C, Winch PJ, Mangwadu G, Arnold BF, Prendergast AJ, et al. The WASH Benefits and SHINE trials: interpretation of WASH intervention effects on linear growth and diarrhoea. Lancet Glob Health. 2019;7(8):e1139-e46.

33. Rego R, Watson S, Alam MAU, Abdullah SA, Yunus M, Alam IT, 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.

Corresponding author: Ryan Rego, Center for Global Health Equity, University of Michigan at Ann Arbor, Ann Arbor, USA. Phone: +1 810 955 0949, Email: RegoR@UMich.edu