The impact of sanitation on infectious disease and nutritional status: A systematic review and meta-analysis

Matthew C. Freeman a,*, Joshua V. Garn b, Gloria D. Sclar a, Sophie Boisson b, Kate Medlicott b, Kelly T. Alexander a, Gauthami Penakalapati a, Darcy Anderson a, Amrita G. Mahtani a, Jack E.T. Grimes c, Eva A. Rehfuess d, Thomas F. Clasen a,*,c

a Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA
b Department of Public Health, Environmental and Social Determinants of Health (PHE), World Health Organization, Geneva, Switzerland
c Department of Civil and Environmental Engineering, South Kensington Campus, Imperial College London, London, UK
d Institute for Medical Informatics, Biometry and Epidemiology, Pettenkofer School of Public Health, LMU Munich, Germany
e Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK

A R T I C L E   I N F O

Article history:
Received 9 March 2017
Received in revised form 16 May 2017
Accepted 18 May 2017

Keywords:
Sanitation
Diarrhea
Soil-transmitted helminths
Trachoma
Schistosomiasis
Nutrition

A B S T R A C T

Background: Sanitation aims to sequester human feces and prevent exposure to fecal pathogens. More than 2.4 billion people worldwide lack access to improved sanitation facilities and almost one billion practice open defecation. We undertook systematic reviews and meta-analyses to compile the most recent evidence on the impact of sanitation on diarrhea, soil-transmitted helminth (STH) infections, trachoma, schistosomiasis, and nutritional status assessed using anthropometry.

Methods and findings: We updated previously published reviews by following their search strategy and eligibility criteria. We searched from the previous review’s end date to December 31, 2015. We conducted meta-analyses to estimate pooled measures of effect using random-effects models and conducted subgroup analyses to assess impact of different levels of sanitation services and to explore sources of heterogeneity. We assessed risk of bias and quality of the evidence from intervention studies using the Liverpool Quality Appraisal Tool (LQAT) and Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach, respectively. A total of 171 studies met the review’s inclusion criteria, including 64 studies not included in the previous reviews. Overall, the evidence suggests that sanitation is protective against diarrhea, active trachoma, some STH infections, schistosomiasis, and height-for-age, with no protective effect for other anthropometric outcomes. The evidence was generally of poor quality, heterogeneity was high, and GRADE scores ranged from very low to high.

Conclusion: This review confirms positive impacts of sanitation on aspects of health. Evidence gaps remain and point to the need for research that rigorously describes sanitation implementation and type of sanitation interventions.

© 2017 World Health Organization Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND IGO license (http://creativecommons.org/licenses/by-nc-nd/3.0/igo/).
1. Introduction

An estimated 2.4 billion people lack access to improved sanitation—pit latrines with slabs or other facilities intended to sequester human feces from the environment (WHO/UNICEF, 2013). Almost one billion of these people have no sanitation facility whatsoever and practice open defecation. Nearly all these sanitation deficiencies are among vulnerable populations in low-income countries, and are primarily in rural settings and urban slums in South and Southeast Asia and Sub-Saharan Africa (WHO/UNICEF, 2015).

Poor sanitation may be associated with a number of infectious and nutritional outcomes, and these outcomes also cause a heavy burden of disease globally. Diarrhea accounts for the largest share, causing an estimated 1.4 million deaths annually (Lozano et al., 2012; Pruss-Ustun et al., 2014) or 19% of all under-five deaths in low-income settings (Boschi-Pinto et al., 2008). Over one billion people are at risk of soil-transmitted helminth (STH) infections, which leads to nearly five million disability adjusted life years (DALYs), or in other words five million years of healthy life lost, while schistosomiasis causes the loss of a further two million (Murray et al., 2013; Pullan et al., 2014). Trachoma is the leading infectious cause of blindness in the world (Resnikoff et al., 2004), responsible for visual impairment of 2.2 million people with a total of 1.2 million irreversibly blind (Pascolini and Mariotti, 2011). Globally, 142 million children are stunted (De Onis et al., 2012). In addition to the direct effects of sanitation on human health, sanitation-related sequelae aggravate poverty and economic development (Guerrant et al., 2013).

Diarrheal pathogens include viruses, bacteria, and protozoans, and are primarily transmitted via human feces, though some also have animal hosts (Wagner and Lanois, 1958). Sanitation is considered a primary barrier to infection by excluding pathogens from the environment (Wagner and Lanois, 1958), though rotavirus, the largest global contributor to diarrheal disease in young children (Kotloff et al., 2013) is not prevented by improved sanitation. Nearly all cases of soil-transmitted helminthiases, schistosomiasis, and trachoma are environmentally mediated (Prüss-Ustün et al., 2016).
and consistent use of sanitation and hygienic behaviors is likely to play a role in preventing transmission.

STHs are parasitic nematodes that live in the gut and are spread primarily through fecal contamination of the environment (Hotez et al., 2006). Over one billion people are at risk of STH infections throughout the world (thiswormyworld.org) as exposure occurs in places where human excreta is not contained or treated, when larvae, after hatching in soil infect human hosts either through the skin (hookworm or Strongyloides stercoralis) or ingestion (Ascaris lumbricoides, Trichuris trichiura, and sometimes Ancyclostoma duodenale). STHs can be treated safely and somewhat effectively with several approved drugs (Keiser and Utzinger, 2008); though there has been some debate on the appropriate approach for delivery of these drugs (Taylor-Robinson et al., 2015). The World Health Organization (WHO) recommends mass treatment among endemic populations (WHO, 2005).

Trachoma is caused by a bacterial infection (Chlamydia trachomatis) that causes scarring of the inner eyelid; repeated infection over a lifetime can lead to inward turning of the eyelid and, without treatment, corneal opacity and blindness (Hu et al., 2010). It is endemic throughout sub-Saharan Africa and Asia, and parts of Latin America (TrachomaAtlas.org). C. trachomatis is spread through scalyptic flies that breed in human feces; reducing the number of fly breeding sites through improvements in sanitation is thought to reduce transmission. Comprehensive control of trachoma is conducted using the “SAFE” strategy (Emerson et al., 2006), which includes Surgery to correct severe cases of active trachoma, Antibiotics distributed to all at-risk community members bi-annually, promotion of Facial hygiene, and Environmental improvements, typically promotion of sanitation.

Schistosomiasis is caused by a parasitic trematode, whose eggs are expelled either in stool (Schistosoma mansoni and S. japonicum) or urine (S. haematobium); eggs hatch to release miracidia, which infect freshwater snails before dividing into free-swimming cercariae that infect the human definitive host through skin contact (Hotez et al., 2006). Activities causing contact with infested water and therefore Schistosoma infection include bathing, farming, water handling, laundry, and recreational swimming. Long term and heavy infections increase the risk of serious adverse health outcomes, including bladder cancer in the case of S. haematobium infections (Hotez et al., 2006). Periodic treatment with a single dose of praziquantel is effective at preventing the development of heavy infections and long-term sequelae, but reinfection generally follows treatment (Engels et al., 2002).

Poor sanitation can adversely impact nutritional status in young children not only through the impaired absorption of nutrients associated but through sub-clinical infections with fecal pathogens (Guerrant et al., 2012; Humphrey, 2009). Repeated and persistent infection may lead to environmental enteric dysfunction, a sub-clinical condition that can lead to growth faltering (Ngure et al., 2014).

Previous systematic reviews found sanitation to be protective against diarrhea (Pruss-Ustun et al., 2014), STH infections (Strunz et al., 2014), trachoma (Stocks et al., 2014), schistosomiasis (Grimes et al., 2014), and poor nutritional status (Dangour et al., 2013). However, the studies included in these reviews were mainly observational or small-scale trials, most of which combined sanitation with water supplies or hygiene. While some of these reviews assess the methodological quality or risk of bias of the included studies, none seek to assess the quality of the overall body of evidence. Moreover, several of the more rigorous trials to assess the impact of sanitation on diarrhea, STH infection, and nutritional outcomes were not included in these prior reviews (Briceno et al., 2015; Clasen et al., 2014a,b; Patil et al., 2014; Pickering et al., 2015). Many of the more recent, rigorous trials have found no effect, or mixed effects for these outcomes, and so we explore in our review the role of sanitation coverage and use across these studies. Because many of the outcomes of this review share transmission mechanisms, there is merit in assessing and reporting on these outcomes together.

As part of its effort to develop guidelines on sanitation and health, the WHO commissioned this systematic review to examine the effect of sanitation on major infectious diseases and nutritional outcomes in populations around the world. For this purpose, we updated several previously published systematic reviews and conducted additional sub-group analyses including assessing the health impact of different levels of sanitation services as defined by the WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation (JMP) (WHO/UNICEF, 2015).

2. Methods

2.1. Search strategy

As this review was designed to update previously published reviews on diarrhea (Pruss-Ustun et al., 2014), STH infection (Strunz et al., 2014), trachoma (Stocks et al., 2014), schistosomiasis (Grimes et al., 2014), and nutritional status (Dangour et al., 2013), we followed the study-specific protocols of the designated reviews (available upon request), including search strategies (Table S1). We conducted separate database searches for each health outcome and searched the same set of databases utilizing the same Boolean search strings as the previous review. However, we employed a time restriction that started from the previous review’s last end point up to December 31, 2015. Searches were conducted in English, though we included studies published in English, Spanish, Portuguese, French, German, or Italian. Eligible studies were included regardless of publication status (published, unpublished, in press, grey literature, etc.). In addition to the electronic searches, we hand-searched references from all included studies. We followed the PRISMA guidelines in this review.

2.2. Study eligibility

2.2.1. Type of populations

We included all age groups and country settings, except in the case of the nutrition review by Dangour et al. (2013) which excluded studies with adults. We included any kind of special sub-groups such as people living with HIV/AIDS and special settings such as schools.

2.2.2. Type of outcome measures

Primary health outcomes were diarrhea, STH infections, trachoma, schistosomiasis, and nutritional status.

2.2.3. Diarrhea

Diarrhea was self-reported, clinically confirmed or other recorded morbidity associated with diarrhea, including acute, persistent, bloody, and watery diarrhea and dysentery. This includes the WHO definition of diarrhea, which is three or more loose stools in 24 h (Baqui et al., 1991). In studies that did not use such a definition, we used the same case definition that was used in the study.

2.2.4. STH infections

We assessed infection with STHs and considered studies that measured the four predominant worms individually – A. lumbricoides, T. trichiura, hookworm (Ancylostoma duodenale and Necator americanus), and S. stercoralis. Due to the different pathways of infection for each worm species, studies that did not specify the worm species were not considered eligible.
2.2.5. Schistosomiasis
Infection with various schistosomes (S. mansoni, S. japonicum, and S. haematobium) were assessed independently.

2.2.6. Trachoma
Trachoma outcomes included active trachoma — presence of trachomatous inflammation-follicular or trachomatous inflammation-intense (TF/TI) — or laboratory confirmed presence of C. trachomatis infection diagnosed using PCR.

2.2.7. Nutritional status
The nutrition outcomes included underweight, wasting, and stunting, which were all measured through anthropometry (weight, age, and height) and based on WHO child growth standards. All study settings and populations were eligible, except for the anthropometric outcomes, which only included children as per the protocol for the nutritional status review by Dangour et al. (2013).

2.2.8. Type of sanitation exposures and comparisons
We reviewed experimental and observational studies that reported on the effect of sanitation, be it the presence or use of sanitation or the effect of specific sanitation interventions, compared to no sanitation or lower levels of sanitation. Eligible sanitation interventions included provision of sanitation facilities or services (e.g., provision of household latrines or child potties) and promotional activities (e.g., behaviour change promotion to reduce open defecation). Interventions that combined the safe disposal of feces with other interventions, such as improvements to water supply, water quality, or promotion of hygiene, were included in the review and subgroup analyses were conducted to investigate their impact on the effect estimate. A more detailed description of the studies can be found in Tables S2–S13.

2.2.9. Type of study designs
Eligible study designs included randomized controlled trials (RCTs) and non-randomized studies (NRS): quasi-RCTs, non-randomized controlled trials, controlled before-and-after studies, interrupted-time-series studies, historically controlled studies, case-control studies, cohort studies (uncontrolled before-and-after studies) and cross-sectional studies. Two reviews — the Dangour et al. (2013) review of nutritional status and the Wolf et al. (2014) review of diarrhea — included further restrictions on the types of observational studies; as such, we followed the same protocols, and the Dangour et al. (2013) update included no cross-sectional studies and the Wolf et al. (2014) update only included cross-sectional studies that used a specific matching method, such as propensity score matching.

2.3. Selection of studies
As previously mentioned, we conducted separate database searches for each health outcome. For each search, one reviewer performed a single screening of clearly non-relevant titles. Next, two reviewers independently reviewed the remaining abstracts to determine if they met the inclusion criteria for the review. If a title or abstract could not be rejected with certainty, the full text was obtained for further screening. We contacted authors of studies when additional data was needed to assess eligibility for inclusion.

After obtaining full copies of all potentially relevant studies, two reviewers scanned the full texts to determine if they met the inclusion criteria. When the two reviewers were unable to agree on inclusion or exclusion, a third reviewer was consulted to make the final decision. Reasons for exclusions of articles excluded based on full text are provided in the PRISMA flow diagram in Fig. 1.

2.4. Data extraction and management
Data for each disease outcome was independently extracted by two reviewers from the newly selected studies using a piloted data extraction form. One reviewer extracted data from the studies included in the previous reviews. Categories of data to be entered into the form are found in the supplemental materials (Spreadsheet S1) and include study characteristics, description of study population, description of intervention and outcome measures, sanitation, water and hygiene characteristics of the study population, and reported or calculated results. For studies with missing data, we attempted to contact authors to supply additional information. Following data extraction, forms were compared, and in case of a discrepancy, consensus was reached. In case of continued disagreement, a third author made the final decision about how to interpret the data.

2.5. Assessment of risk of bias
We used an abridged version of the Liverpool Quality Appraisal Tool (LQAT) to assess risk of bias because of the tool’s flexibility in accommodating different study designs and in creating exposure and outcome assessments specific to our diverse set of sanitation and health studies (Pope et al., 2010; Puzzolo et al., 2016). Due to the large number of experimental studies and the greater degree of bias inherent in observational study designs, we only assessed bias in our experimental studies (see also section on assessment of quality of evidence below). The LQAT considers eight areas of bias: selection bias, response rate bias, allocation bias, follow-up bias, bias in exposure assessment, bias in outcome assessment, bias in ascertainment, and confounding in analysis. An average LQAT score of nine or above indicated relatively low risk of bias while a score of five to eight indicated serious risk of bias and a score of four or below very serious risk of bias. Two reviewers independently assessed risk of bias while a third reviewer compared the scores and resolved any discrepancies. The LQAT can be found in the supplemental materials (Spreadsheet S2).

2.6. Data synthesis
Where studies reported both adjusted and unadjusted results, we used the most adjusted estimate assessing the effect of sanitation on the outcome; we also used the corresponding 95% confidence interval (CI) reported by the study. If adjusted estimates were unavailable, we then used the unadjusted measures of effect reported by the study or calculated an unadjusted estimate and 95% CI based on raw data and reported p-values. As we were updating previous reviews, we used the same measure of effect (e.g. odds ratio, risk difference) as was used in the previous review. When a study utilized a different measure of effect than the one chosen for the meta-analysis, raw data reported in the study was used to convert the effect into the appropriate point estimate and 95% CI. If raw data was not available, however, we followed the guidance from the Cochrane Collaboration on how to convert effect estimates (Higgins et al., 2008). In studies that reported multiple sanitation comparisons and thus multiple measures of effect, we used the raw data reported in the study to calculate weighted unadjusted measures of effect in order to prevent the counterfactual group from being counted more than once in the meta-analysis. If the necessary raw data was not reported in the study, then only one measure of effect was extracted. In these cases, we resorted to the more unambiguous sanitation comparison or when authors reported comparisons for access to a latrine and use of a latrine, we deferred to the use comparison as a better measure for assessing the impact of sanitation. Similarly, if a study reported measures of effect for sanitation on different age groups, the youngest of the age groups was selected
for the meta-analyses. This did not apply for the sub-group analyses that stratified by age (under-two years of age, under-five years of age, school-aged children, adults) where all measures of effect were included.

We conducted meta-analyses to estimate pooled measures of effect. Due to the large heterogeneity in populations, interventions and outcome measures and as a result the expected large statistical heterogeneity, we decided a priori to use random-effects models.

For each disease outcome, our analysis included four separate comparisons:

1. All studies. We compared the overall impact of sanitation by pooling the primary effect estimates of studies that met our inclusion criteria. We show these results in forest plots, where the point estimates and confidence intervals are shown for each study, and a diamond (and dotted vertical line) represents the overall

---

**Fig. 1.** PRISMA flow diagram of publications considered for the review. The order for studies is total, diarrhea, STH, trachoma, schistosomiasis, and nutrition.
summary estimate with the width of the diamond representing the confidence interval. These primary point estimates could have been assessing one of several different types of sanitation comparisons: comparisons of access to any sanitation vs. no sanitation, access to improved sanitation vs. unimproved sanitation (as defined by the JMP (WHO/UNICEF, 2008)) or those within sanitation intervention arm vs. those in the control arm. Where possible we estimated an effect for comparisons of sanitation “use” vs. “non-use.” However, few studies explicitly measured sanitation use as the vast majority only assessed access to or presence of sanitation.

2. Intervention studies. Since many of our included studies were observational, we isolated the experimental studies that specifically assessed a sanitation intervention in order to provide a more rigorous pooled estimate. We also discuss each of the more rigorous intervention studies individually. We also assess how the health outcomes are associated with latrine coverage and latrine use levels (Garn et al., 2016) for all of the intervention studies where available that included information on both health and coverage/use, to better characterize latrine coverage and latrine use thresholds for improving health (Table S15).

3. Sanitation ladder. We wanted to assess the impact on health for different types of sanitation by pooling estimates for several comparisons along the sanitation ladder, a term defined by JMP (WHO/UNICEF, 2008) that defines the level of service as incremental improvement in sanitation conditions from open defecation (no sanitation) to unimproved sanitation to shared sanitation and finally, to improved sanitation (WHO/UNICEF, 2016). We calculated estimates comparing any sanitation facility to none and improved sanitation to both unimproved and shared sanitation facilities. Only a small number of studies provided enough detail on the type of sanitation assessed for us to determine a sanitation condition category based on JMP definitions.

4. Stratified analyses. We explored additional characteristics of study populations and sub-populations such as: study setting (rural vs. urban), age group, geographic region, season, length of follow-up post intervention implementation (only for intervention studies), and water and soap coverage among the study population. Water coverage was defined by the percent of the study population with access to an improved drinking water source and soap coverage was defined by the percent of the study population practicing handwashing with soap or other cleaning agent or was observed to have soap present somewhere in the household compound.

When there were an insufficient number of similar studies to pool, the individual study results were not pooled and a description of the results was presented in tabular form. Data were analysed using STATA 14 (StataCorp; College Station, TX, USA).

2.7. Heterogeneity

Study heterogeneity was examined in relation to populations, interventions, settings, outcome measures as well as study designs. Statistical heterogeneity was assessed by visually examining the confidence intervals in the forest plot and by using the $I^2$ statistics. We considered an estimate of $I^2 > 50\%$ to indicate there may be substantial heterogeneity and $I^2 > 75\%$ to indicate there may be considerable heterogeneity (Higgins et al., 2008). The Cochrane handbook advises that the importance of $I^2$ depends on the magnitude and direction of effects and also the evidence for heterogeneity (Higgins et al., 2008). While some have argued against pooling with high heterogeneity, Caldwell and Welton (2016) have argued that pooling may still be appropriate for general effectiveness questions. We followed the protocols of previous reviews, pooling estimates, as they did in their reviews. However, we also present results from every study in supplemental materials, and discuss in the text each of the more rigorous intervention studies individually in hopes of presenting both viewpoints.

2.8. Assessment of the quality of evidence

The GRADE approach was used to assess the quality of the evidence from intervention studies for each health outcome (Schünemann et al., 2013). GRADE scores range from high quality of evidence to moderate to low to very low and can be interpreted as the level of confidence one has that the estimated effect from a given body of evidence is close to the true effect (Guyatt et al., 2011). Given the relatively large number of intervention studies, we focused our GRADE assessment on these studies: observational studies were used as supporting evidence. In the GRADE approach, RCTs start out with a ‘high’ score while observational studies start at a ‘low’ score. For this review, since some of the intervention studies were RCTs and all had a control comparison group, we started the body of evidence at a ‘high’ score. Besides the design of the underlying studies, the GRADE approach uses a separate set of five criteria to further determine whether to downgrade the quality of the evidence – risk of bias of individual studies, inconsistency, indirectness, imprecision and publication bias (Guyatt et al., 2011). Each body of evidence could be downgraded up to two points for each criterion depending on whether the issue was serious (-1) or very serious (-2). Risk of bias was assessed by using the averaged LQAT score of the combined intervention studies and downgrading for serious (-1) and very serious (-2) risk of bias. Inconsistency was primarily assessed by whether or not there was a difference in the direction of effect of the point estimates. However, statistical heterogeneity was also examined through the $I^2$ and Chi²-p value statistical tests of the pooled effect estimate. Indirectness refers to how directly the evidence addressed the review’s research question of interest in regards to the intervention or exposure of interest, population, and measured outcome (Guyatt et al., 2011). However, as this review considered any population, geographical setting and a range of sanitation interventions, we did not downgrade for indirectness since the included studies all addressed this broader research question and at least one of the outcomes of interest.

We downgraded the evidence for lack of precision if the pooled effect estimate’s CI overlapped with the null (i.e. one for ratio estimates, zero for difference estimates) and was thus not statistically significant. Publication bias was assessed by visually examining the level of symmetry in the corresponding funnel plot.

3. Role of the funding source

Funding for this article was provided to by The Bill & Melinda Gates Foundation and UK Aid to the WHO. Some authors from the WHO were involved in the study design and provided feedback on this manuscript. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

4. Results

4.1. Eligible studies

We identified a total of 12,106 articles during the database search process for all health outcomes with 10,485 articles remaining after deletion of duplicates (Fig. 1). Nearly all were from electronic database searches (11,989). After reviewing the titles and abstracts of 10,485 publications, it was determined that 9,464 articles did not meet the inclusion criteria and 1,022 articles were
Table 1: Association with and impact of sanitation on diarrhea, STH, trachoma, schistosomiasis, and nutrition and quality of evidence.

<table>
<thead>
<tr>
<th></th>
<th>Overall sanitation comparison*</th>
<th>Intervention vs. control#</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Pooled ORs</td>
</tr>
<tr>
<td>Diarrhea</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. lumbricoides</td>
<td>38</td>
<td>0.73 (0.61, 0.86)</td>
</tr>
<tr>
<td>T. trichiura</td>
<td>33</td>
<td>0.80 (0.63, 1.03)</td>
</tr>
<tr>
<td>Hookworm</td>
<td>40</td>
<td>0.65 (0.53, 0.78)</td>
</tr>
<tr>
<td>S. stercoralis</td>
<td>9</td>
<td>0.48 (0.36, 0.65)</td>
</tr>
<tr>
<td>Trachoma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. pallidum</td>
<td>39</td>
<td>0.70 (0.62, 0.79)</td>
</tr>
<tr>
<td>C. trachomatis</td>
<td>10</td>
<td>0.62 (0.44, 0.87)</td>
</tr>
<tr>
<td>Schistosomiasis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. mansoni</td>
<td>23</td>
<td>0.61 (0.50, 0.74)</td>
</tr>
<tr>
<td>S. haematobium</td>
<td>10</td>
<td>0.69 (0.58, 0.81)</td>
</tr>
<tr>
<td>Nutrition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight-for-age Z</td>
<td>7</td>
<td>0.02 (−0.05, 0.08)</td>
</tr>
<tr>
<td>Weight-for-height Z</td>
<td>3</td>
<td>−0.03 (−0.11, 0.04)</td>
</tr>
<tr>
<td>Height-for-Z</td>
<td>9</td>
<td>0.08 (0.00, 0.16)</td>
</tr>
</tbody>
</table>

OR = odds ratio; DF = mean difference; NA = not available.

* The overall sanitation comparison included any sanitation/intervention/improved vs. none/control/unimproved. Studies comparing improved vs. shared sanitation were not included because we didn’t deem the “shared” group to be a comparable control group.

# The intervention vs. control comparison includes trials and other studies that assessed the impact of a sanitation intervention on diarrhea.

? We followed the protocol from Dangour et al., which excluded cross-sectional and case-control studies; as such we only present intervention studies.

? Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach (Schünemann et al., 2013).

4.2. Diarrhea

Overall, 33 studies met our eligibility criteria (Fig S1) and 27 were included in meta-analyses (Table 1). These include 11 studies from the previous review by Wolf et al. (2014) and 16 intervention studies. Studies were conducted in Asia (N = 11), Africa (N = 8), Central and South America (N = 5), and the Eastern Mediterranean (N = 3) (Table S2). Twenty-two of the studies assessed impact among children under five years.

4.2.1. Overall effect of sanitation

Overall, sanitation was associated with 12% lower odds of diarrhea (OR 0.88, 95% CI 0.83–0.92, N = 27) (Table 1; Fig 2), though this pooled estimate is characterized by high heterogeneity (I² = 79.0%).

4.2.2. Intervention studies

When restricted to the 16 intervention studies, sanitation interventions resulted in a pooled OR estimate of 0.77 (95% CI 0.66–0.91) (Table 1). Eight randomized controlled trials met the inclusion criteria. Three assessed latrine provision interventions implemented under the “Total Sanitation Campaign” in India and reported no effect on diarrhea (Clasen et al., 2014a,b; Dickinson et al., 2015; Patil et al., 2014) (Fig 3). In Mali, a community-led total sanitation (CLTS) intervention also found no impact on diarrhea (Pickering et al., 2015). In Tanzania a CLTSlike intervention that also included sanitation marketing, and promoting and enabling environment found no decrease in diarrhea (Briceño et al., 2015). However, a large-scale CLTS program in Indonesia reported a 49% reduction in diarrhea prevalence despite similar sanitation uptake between study arms (Cameron et al., 2013). A hygiene education intervention with safe disposal messaging found lower diarrhea in the intervention compared to the control arms (Stanton et al., 1988). A school-based cluster-randomized trial of a comprehensive water, sanitation, and hygiene (WASH) intervention found no effect on diarrhea among school-age children (Freeman et al., 2013b) and their under-five siblings (Dreibelbis et al., 2014) in a subgroup of schools that already had access to water, but found reduced diarrhea among the sub-group of schools that were from water-scarce schools that also received water supply as part of the intervention. There were also other rigorous interventions using non-randomized designs (See Supplemental Table 2). Several of intervention studies reported measures of effect other than the odds ratios, and so were not able to include these in the meta-analyses (Arnold et al., 2010; Barreto et al., 2007; Briceño et al., 2015; Kariuki et al., 2012; Stanton et al., 1988) (Fig 4). Of the householdbased diarrhea intervention studies, only 9 reported on either latrine coverage or latrine use (Table S14) (Arnold et al., 2010; Barreto et al., 2007; Briceño et al., 2015; Cameron et al., 2013; Clasen et al., 2014a,b; Moraes et al., 2003; Patil et al., 2014; Pickering et al., 2015; Pradhan and Rawlings, 2002). Only three of these nine studies found a decrease in diarrhea due to the intervention (Barreto et al., 2007; Cameron et al., 2013; Moraes et al., 2003); two were sewerage studies and the other did not even find an increase in latrine coverage and attributed the health gains to probably being due to drinking water and handwashing behavior. Eight of these nine diarrhea studies reported increases in latrine coverage (i.e. increases between 6% and 51%; Table S14). In those papers where no effect of coverage on diarrhea was observed, authors explained their results in different ways.

4.2.3. GRADE for intervention studies

The quality of the evidence from the 16 intervention studies reporting on the impact of sanitation on diarrhea was scored ‘low’...
4.2.4. Sanitation ladder

The pooled odds ratio from four studies that compared those that used or had access to any sanitation facility with those who did not was 0.76 (95% CI 0.64–0.91) (Table 2). Pooled estimates from seven studies show a 4% reduction in the odds of diarrhea from improved versus unimproved sanitation (OR 0.96, 95% CI 0.92–0.99). Four studies assessed improved versus shared sanitation (OR 0.48, 95% CI 0.20–1.16). There may be other papers that assess the sanitation ladder, but our review used the same search criteria as Wolf et al. (2014), which only included certain types of observational studies (e.g., those using specific matching methods) (Fig. 6).

4.2.5. Subgroup analyses

Sanitation was protective against diarrhea among children under-five (OR 0.91, 95% CI 0.86–0.95, N = 22) and among study populations where age groups were not specified (OR 0.61, 95% CI 0.47–0.79, N = 2) (Table S16). The effect estimate was similar but not statistically significant among school-aged children (OR 0.71, 95% CI 0.48–1.04, N = 3), and no effect was found among adults alone (OR 1.03, 95% CI 0.91–1.16, N = 3). Studies conducted in rural settings found a significant protective effect on diarrhea (OR 0.80, 95% CI 0.66–0.96, N = 11), while those conducted in urban settings...
found none (OR 0.87, 95% CI 0.53–1.43, N = 4). The impact of sanita-
tion on diarrhea was greater in areas with higher levels of access
to improved water supply while soap coverage showed no effect
on the impact of sanitation. Estimates of effect of sanitation by
length of follow-up showed lower odds of diarrhea in sanitation
studies with less than six months of follow-up, and also in sanita-
tion studies with more than twelve months of follow-up. The level
of soap coverage in the study population did not seem to influence
the impact of sanitation on diarrhea (Fig. 7).

4.3. Soil-transmitted helminth infections

Overall, 65 studies met our eligibility criteria, including 31
from the previous review by Strunz et al. (2014) (Fig. S2). These

![Fig. 3](image-url) Meta-analysis examining the effectiveness of sanitation interventions on *A. lumbricoides* prevalence, as well as comparing the association between different rungs of the sanitation ladder and *A. lumbricoides* prevalence.

---

**Table: Sanitation Interventions**

<table>
<thead>
<tr>
<th>Intervention vs. control</th>
<th>Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barreto, 2010 (CLTS location B - 36 month follow up)</td>
<td>0.42 (0.33, 0.54)</td>
</tr>
<tr>
<td>Beltrán, 2015 (CLTS location A - 15 month follow up)</td>
<td>0.42 (0.12, 1.47)</td>
</tr>
<tr>
<td>Cameron, 2013</td>
<td>1.03 (0.63, 1.68)</td>
</tr>
<tr>
<td>Classen, 2014</td>
<td>2.48 (1.32, 4.64)</td>
</tr>
<tr>
<td>Freeman, 2013</td>
<td>0.56 (0.31, 1.00)</td>
</tr>
<tr>
<td>Gelaye, 2014</td>
<td>0.50 (0.24, 1.03)</td>
</tr>
<tr>
<td>Gyorkos, 2013</td>
<td>0.88 (0.57, 1.34)</td>
</tr>
<tr>
<td>Henry, 1988</td>
<td>0.46 (0.30, 0.71)</td>
</tr>
<tr>
<td>Pali, 2014</td>
<td>0.89 (0.56, 1.47)</td>
</tr>
<tr>
<td>Steinmann, 2015</td>
<td>0.24 (0.13, 0.44)</td>
</tr>
<tr>
<td>Subtotal (I-squared = 88.0%, p = 0.000)</td>
<td>0.77 (0.49, 1.21)</td>
</tr>
</tbody>
</table>

---

**Table: Improved vs. unimproved**

<table>
<thead>
<tr>
<th>Any sanitation vs. none/non-use</th>
<th>Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Av-Makkial, 2007</td>
<td>0.50 (0.27, 0.91)</td>
</tr>
<tr>
<td>Assou, 2007</td>
<td>0.70 (0.40, 1.15)</td>
</tr>
<tr>
<td>Belo, 2005</td>
<td>1.59 (0.74, 3.45)</td>
</tr>
<tr>
<td>Corrales, 2006 (double-vault desiccating vs. none)</td>
<td>9.77 (1.28, 74.50)</td>
</tr>
<tr>
<td>Corrales, 2006 (pit latrine vs. none)</td>
<td>0.72 (0.37, 1.39)</td>
</tr>
<tr>
<td>Corrales, 2006 (solar latrine vs. none)</td>
<td>0.42 (0.03, 6.97)</td>
</tr>
<tr>
<td>Freeman, 2015</td>
<td>0.83 (0.60, 1.14)</td>
</tr>
<tr>
<td>Greenuk, 2015</td>
<td>0.90 (0.70, 1.10)</td>
</tr>
<tr>
<td>Habbari, 2000</td>
<td>0.71 (0.49, 1.04)</td>
</tr>
<tr>
<td>Kopp, 2013</td>
<td>0.64 (0.40, 1.03)</td>
</tr>
<tr>
<td>Kournavong, 2011</td>
<td>1.08 (0.57, 2.05)</td>
</tr>
<tr>
<td>Morales-Espinosa, 2003</td>
<td>0.62 (0.28, 1.40)</td>
</tr>
<tr>
<td>Mufioz-Antoli, 2014</td>
<td>1.00 (0.40, 2.60)</td>
</tr>
<tr>
<td>Nasr, 2013</td>
<td>0.56 (0.39, 0.81)</td>
</tr>
<tr>
<td>Nguyen, 2005</td>
<td>1.07 (0.93, 1.23)</td>
</tr>
<tr>
<td>Norhayati, 1998</td>
<td>0.53 (0.32, 0.99)</td>
</tr>
<tr>
<td>Olsen, 2001</td>
<td>0.80 (0.45, 1.41)</td>
</tr>
<tr>
<td>Pham-Duc, 2013 (dry latrine vs. no latrine)</td>
<td>1.32 (0.62, 2.79)</td>
</tr>
<tr>
<td>Pham-Duc, 2013 (water-flush latrine vs. dry latrine)</td>
<td>0.92 (0.42, 1.99)</td>
</tr>
<tr>
<td>Phongla, 2013</td>
<td>1.18 (0.59, 2.38)</td>
</tr>
<tr>
<td>Ríosquez, 2010</td>
<td>0.07 (0.01, 0.66)</td>
</tr>
<tr>
<td>Schüle, 2014</td>
<td>0.68 (0.32, 1.43)</td>
</tr>
<tr>
<td>Trang, 2006</td>
<td>0.29 (0.10, 0.86)</td>
</tr>
<tr>
<td>Trong, 2007</td>
<td>0.38 (0.12, 1.21)</td>
</tr>
<tr>
<td>Ugborneo, 2009</td>
<td>0.00 (0.00, 0.01)</td>
</tr>
<tr>
<td>Xiao, 2015</td>
<td>0.51 (0.27, 0.96)</td>
</tr>
<tr>
<td>Young, 2007</td>
<td>0.33 (0.13, 0.82)</td>
</tr>
<tr>
<td>Subtotal (I-squared = 81.4%, p = 0.000)</td>
<td>0.63 (0.49, 0.81)</td>
</tr>
<tr>
<td>Overall (I-squared = 84.7%, p = 0.000)</td>
<td>0.73 (0.51, 0.86)</td>
</tr>
</tbody>
</table>
4.3.1. Overall effect of sanitation

Sanitation was associated with lower odds of infection with *A. lumbricoides* (OR 0.73, 95% CI 0.61–0.86, N = 39), *T. trichiura* (OR 0.80, 95% CI 0.63–1.03, N = 34), hookworm (OR 0.65, 95% CI 0.53–0.78, N = 42), and *S. stercoralis* (OR 0.48, 95% CI 0.36–0.65, N = 9). However, heterogeneity was high for all meta-analyses (Table 1). There was evidence of a protective association between any sanitation (presence or use) compared to no sanitation and for STH worm species including *A. lumbricoides* (OR 0.63, 95% CI 0.49–0.81, N = 24), *T. trichiura* (OR 0.72, 95% CI 0.55–0.95, N = 21), hookworm (OR 0.61, 95% CI 0.51–0.72, N = 30), *S. stercoralis* (OR 0.50, 95% CI 0.36–0.70, N = 7) (Table 2) (Fig. 9).

4.3.2. Intervention studies

When restricted to intervention studies, prevalence of infection was similar in both the intervention and control arms for *A. lumbricoides* (OR 0.77, 95% CI 0.49–1.21, N = 10), *T. trichiura* (OR 1.00, 95% CI 0.37–2.69, N = 8), and hookworm (OR 0.94, 95% CI 0.76–1.15, N = 7) (Table 1). Several of the sanitation intervention studies did find an impact on STH infection when complemented with preventive chemotherapy (PC) programs. A school-based trial of toilet provision and hygiene education in Kenya found substantial protective effects of the intervention on *A. lumbricoides*, but not other STHs (Freeman et al., 2013a). Similarly, in Peru, a sanitation and hygiene education program found a 12% reduction in the odds of *A. lumbricoides* infection among children (Gyorkos et al., 2013). An RCT of participatory hygiene and sanitation educations (PHAST) in Uganda found a decline in STH infections for children under-five that lived in communities that had not received the intervention when compared to those that did, but the finding was not statistically significant (Dumba et al., 2013). In Indonesia, no impact on *A. lumbricoides*, *T. trichiura*, or hookworm was observed in a large-scale sanitation intervention, though it was not clear if PC was conducted in the area (Cameron et al., 2013) (Fig. 10).
Fig. 5. Meta-analysis examining the effectiveness of sanitation interventions on hookworm prevalence, as well as comparing the association between different rungs of the sanitation ladder and hookworm prevalence.

A large-scale sanitation program in urban Salvador, Brazil reduced the prevalence of *A. lumbricoides* infection from 24.4% to 12.0% and *T. trichura* from 18.0% to 5.0%, measured using a before and after trial controlling for a number of key ecological factors (Barreto et al., 2010); it was not clear if the study was carried out in the context of deworming. In China, a five-year study assessing six-month MDA and toilet construction found greater reductions in prevalence of *T. trichiura* and *A. lumbricoides* infection, but not for hookworm or *S. stercoralis*, when compared to PC alone (Steinmann et al., 2015). Very few of the STH intervention studies assessing reported the endline sanitation coverage levels (Table S14) (Fig. 11).

### 4.3.3. GRADE for intervention studies

The quality of the evidence for the impact of sanitation on *A. lumbricoides* and *T. trichiura* was considered ‘very low’ (Table S15). The average LQAT scores were 7.4 and 7.8, respectively, indicating serious risk of bias. The forest plots for both worms also showcased unexplained inconsistency and a lack of precision for the pooled estimate of effect. Hookworm intervention studies also had an average LQAT score of 7.9 and the pooled estimate lacked precision. However, the studies were given a GRADE of ‘low’ quality.
Table 2
Association with and impact of sanitation ladder on diarrhea, STH, trachoma, and schistosomiasis.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Association with and impact of sanitation ladder on diarrhea, STH, trachoma, and schistosomiasis.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Any sanitation vs. none/non-use*</td>
</tr>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>STH</td>
</tr>
<tr>
<td>A. lumbricoides</td>
<td>24</td>
</tr>
<tr>
<td>T. trichiura</td>
<td>21</td>
</tr>
<tr>
<td>Hookworm</td>
<td>30</td>
</tr>
<tr>
<td>S. stercoralis</td>
<td>7</td>
</tr>
<tr>
<td>Trachoma</td>
<td>T/L</td>
</tr>
<tr>
<td>C. trachomatis</td>
<td>34</td>
</tr>
<tr>
<td>S. mansoni</td>
<td>8</td>
</tr>
<tr>
<td>S. haematobium</td>
<td>9</td>
</tr>
</tbody>
</table>

OR = odds ratio. NA = not available.

*No studies found comparing open defecation to unimproved sanitation.

Fig. 6. Meta-analysis examining the effectiveness of sanitation interventions on S. stercoralis prevalence, as well as comparing the association between different rungs of the sanitation ladder and S. stercoralis prevalence.

of evidence due to no unexplained inconsistency. Only one study assessed S. stercoralis and as such, the GRADE criteria were not applied but the study received a LQAT score of 5 suggesting serious risk of bias (Fig. 12).

4.3.4. Sanitation ladder

No evidence was found for the associations with improved sanitation compared to unimproved sanitation for any of the STHs under study. Estimates comparing improved versus shared sanitation were only available for hookworm (OR 0.56 95% CI 0.30–1.03, N = 1) and for S. stercoralis (OR 0.93, 95% CI 0.48–1.81, N = 1).

4.3.5. Subgroup analyses

Most of the A. lumbricoides studies assessed the impact of sanitation on school age children (OR 0.64, 95% CI 0.48–0.85, N = 17; Table S17), given that this is generally the population with the highest worm burden; studies from other age groups did not reveal an association. Sub-grouping for associations between sanitation and T. trichiura only revealed a significant association for studies that assessed all ages (OR 0.76, 95% CI 0.58–0.99, N = 15; Table S18). Most
of the hookworm studies assessed all ages together and the pooled estimate showed a protective association between sanitation and all ages (OR 0.64, 95% CI 0.49–0.83, N = 19; Table S19), school-aged children (OR 0.74, 95% CI 0.57–0.97, N = 13) and adults (OR 0.55, 95% CI 0.42–0.71, N = 5), but not children under the age of five although there were few studies in this group (OR 0.79, 95% CI 0.41–1.52, N = 2). There were too few studies of S. stercoralis for meaningful subgroup analysis (Table S20). Across all four STH outcomes, most studies did not indicate the length of follow-up, seasonality, improved water supply coverage, or soap coverage (Fig. 13).

4.4. Trachoma

Overall, 46 studies met the inclusion criteria and were able to be included in the meta analyses, including 40 previously identified from Stocks et al. (2014) (Fig. S3). Of these 46 studies, 41 studies measured the association between any sanitation and active trachoma (TI/TF), four of which were intervention studies. Ten studies measured the association between any sanitation and C. trachomatis infection, two of which were intervention studies. Nearly all studies (N = 42) were from Sub-Saharan Africa and from rural (N = 31) contexts (Tables S7 and S8).

4.4.1. Overall effect of sanitation
Sanitation was associated with lower odds of active trachoma, measured through assessment of clinical signs of TI/TF (OR 0.70, 95% CI 0.62–0.79, N = 39) and prevalence of lab-validated C. trachomatis infection (OR 0.62, 95% CI 0.44–0.87, N = 10) (Table 1).
4.4.2. Intervention studies
A meta-analysis of intervention studies revealed a protective effect of sanitation on active trachoma (OR 0.51, 95% CI 0.29–0.90, N = 4); only two intervention studies measured C. trachomatis infection (OR 1.02 95% CI 0.76–1.38) (Table 1). Several randomized trials were identified, though none found that the sanitation intervention was able to reduce the odds of active trachoma or C. trachomatis. In one village in Vietnam, nearly 60% of the decline in active trachoma was attributable to community-led water and sanitation improvements as part of the complete SAFE strategy, compared to a village that received surgery and antibiotics alone (Khandekar et al., 2006). Provision of sanitation in The Gambia resulted in similar odds of active trachoma, compared to controls (OR 0.80, 95% CI 0.60–1.07) (Emerson et al., 2004). Provision of a water point and a 3-month “modest” health education program that included a discussion of environmental sanitation (though not sanitation promotion specifically) in five randomly selected villages in Nigeria found no impact on C. trachomatis infection compared to five control villages or active trachoma, though children in both arms were provided tetracycline ointment if active trachoma was observed (Abdou et al., 2010). In a randomized trial, after twenty-four months of follow-up, active trachoma and prevalence of C. trachomatis in communities that received azithromycin at baseline and intensive latrine promotion was no different than in areas that received treatment only (OR 0.90, 95% CI 0.71–1.14) (Stoller et al., 2011). In one of the few sanitation interventions to report on mortality, this trial found no additional benefit of latrine promotion on mortality beyond the effect of antibiotics in trachoma endemic communities of Ethiopia (Gebre et al., 2011).

4.4.3. GRADE for intervention studies
The body of evidence linking sanitation interventions to active trachoma was assessed as ‘high’ quality while the evidence for C. trachomatis was assessed as ‘moderate’ quality. The active trachoma intervention studies had an average LQAT score of 8.5 suggesting low risk of bias, the pooled estimate was statistically significant and the funnel plot symmetrical. While there was considerable statistical heterogeneity, all point estimates had the same direction of effect indicating sanitation to be protective and only varied in effect size. Although the C. trachomatis intervention studies also had an average LQAT score of 10.5 suggesting low risk of bias, the pooled estimate of effect was not statistically significant.

4.4.4. Sanitation ladder
The pooled estimates measuring any sanitation compared to no sanitation revealed an association with active trachoma (OR 0.73, 95% CI 0.64–0.83, N = 34) and C. trachomatis (OR 0.54, 95% CI 0.36–0.81, N = 8) (Table 2). A similar association was found between improved and unimproved sanitation for active trachoma (OR 0.62, 95% CI 0.55–0.70, N = 1), though we found no studies assessing this relationship for C. trachomatis. No association was found comparing improved sanitation to shared sanitation for active trachoma (OR 1.03, 95% CI 0.83–1.28, N = 4) and C. trachomatis (OR 1.03, 95% CI 0.76–1.39, N = 1).

4.4.5. Subgroup analyses
In our assessment of sub-populations, we found associations between sanitation and active trachoma among children under five years of age (OR 0.62, 95% CI 0.43–0.90, N = 8), school-age children...
Fig. 9. Meta-analysis examining the effectiveness of sanitation interventions on \textit{S. mansoni}, as well as comparing the association between different rungs of the sanitation ladder and \textit{S. mansoni}.

(OR $0.71$, 95% CI $0.61$–$0.83$, N = 24), and for estimates where age groups were not specified (OR $0.66$, 95% CI $0.49$–$0.89$, N = 7) (Table S21). As indicated, nearly all these studies were conducted in rural areas (N = 26). Most studies did not report the length of follow-up, seasonality, population improved water supply coverage, or soap coverage. Few studies were available to assess subgroup analysis for associations between sanitation and \textit{C. trachomatis} (Table S22).

4.5. \textbf{Schistosomiasis}

Of 30 studies that met our eligibility criteria for sanitation and schistosomiasis (\textit{S. mansoni}: N = 23; \textit{S. haematobium}: N = 10) (Fig. S4), 22 were identified from Grimes et al. (2014) (Fig. S4); there were no intervention studies for any of the schistosomiasis outcomes (Table 1). Studies were conducted in Sub-Saharan Africa (N = 13) and South America (N = 12), predominantly in Brazil for \textit{S. mansoni}. Most studies were conducted in rural contexts (N = 19; Tables S9 and S10).

4.5.1. \textbf{Overall effect of sanitation}

Sanitation was associated with lower odds of \textit{S. mansoni} (OR $0.61$, 95% CI $0.50$–$0.74$, N = 23) and \textit{S. haematobium} (OR $0.69$, 95% CI $0.58$–$0.81$, N = 10); however, heterogeneity was high for these studies (Table 1). Only one study by Yang et al. (2015) set in rural China looked at the association of household sanitation vs. lack of household sanitation with odds of \textit{S. japonicum} infection, and it did not reveal a statistically significant difference (OR $0.77$, 95% CI $0.40$, 1.43).

4.5.2. \textbf{Intervention studies and GRADE}

We did not find any intervention studies assessing sanitation and schistosomiasis, and so no GRADE score was calculated.
### 4.5.3. Sanitation ladder

For pooled estimates measuring any sanitation compared to no sanitation, we found associations with *S. mansoni* (OR 0.57, 95% CI 0.47–0.70, N = 20) and *S. haematobium* (OR 0.69, 95% CI 0.58–0.81, N = 9) (Table 2). We found no association between improved and unimproved sanitation with *S. mansoni* (OR 1.16, 95% CI 0.49–2.75, N = 3) and *S. haematobium* (OR 0.75, 95% CI 0.24–2.42, N = 1). We found no studies assessing improved versus shared sanitation.

### 4.5.4. Subgroup analyses

In our assessment of sub-populations for *S. mansoni*, studies found preventive associations whether they were conducted on school age children (OR 0.47, 95% CI 0.36–0.62, N = 10) or on participants where age was not specified (OR 0.70, 95% CI 0.53–0.92, N = 12) (Table S23). The association between sanitation and *S. mansoni* was similar in both urban (OR 0.79, 95% CI 0.53–1.16, N = 7) and rural contexts (OR 0.60, 95% CI 0.45–0.79, N = 11). Pooled estimates in sites with both high and low access to improved water showed protective associations with sanitation. The estimates of association were similar in areas with both low and high coverage with improved water supply. There were too few studies to meaningfully assess other subgroup analyses for *S. mansoni* and *S. haematobium* (Table S24).

### 4.6. Nutritional status

Overall, 17 studies met the inclusion criteria, and of these, 14 assessed weight-for-age Z score (WAZ)/underweight, 7 assessed weight-for-height Z scores (WHZ)/wasting, and 14 assessed height-for-age Z scores (HAZ)/stunting. All of these were intervention studies, as we used the same inclusion criteria as Dangour et al. (2013) (Fig. S5). However, not all of these studies assessed our primary Z score outcomes, and were able to be included in the meta-analyses. Seven studies assessed associations between sanitation and weight-for-age Z scores, three studies assessed associations between sanitation and weight-for-age Z scores, and nine studies assessed the associations between sanitation and height-for-age Z scores (Table 1). Studies were primarily conducted in Sub-Saharan Africa for these outcomes (Tables S11-S13).

#### 4.6.1. Intervention studies

Interventions to improve sanitation were not associated with WAZ (Mean difference (MD) 0.02, 95% CI −0.05, 0.08, N = 7) or WHZ (MD −0.03, 95% CI −0.11, 0.04, N = 3). Sanitation was borderline associated with HAZ (MD 0.08, 95% CI 0.00, 0.16, N = 9), and heterogeneity was high this outcome. There were several randomized controlled studies. In Mali, children in communities that randomly received sanitation promotion through community-led
Fig. 11. Meta-analysis examining the effectiveness of sanitation interventions on weight-for-height Z score.

Fig. 12. Meta-analysis examining the effectiveness of sanitation interventions on weight-for-age Z score.
total sanitation (CLTS) had higher height-for-age Z scores (and were less stunted) and higher weight-for-age Z scores (and were less underweight) than those in the control community (Pickering et al., 2015). Another CLTS like intervention in Tanzania found no increases in weight-for-age Z scores, weight for height Z scores, or height-for-age Z scores; lower weight-for-height was even observed in one of the two sanitation arms (Briceño et al., 2015). In Indonesia, an at-scale CLTS program did not find any impact on several growth outcomes (Cameron et al., 2013). In two separate randomized trials of community sanitation promotion resulted in no evidence of improved child growth measures in Madhya Pradesh (Patil et al., 2014) nor Odisha, India (Clasen et al., 2014a,b), though evaluation of a CLTS program in Odisha did reveal an impact on mid-upper-arm circumference, height, and weight (Dickinson et al., 2015). Several intervention studies reported various measures of effect other than the anthropometric Z score differences, and so were not able to include these in the meta-analyses (Ahmed et al., 1993; Belizarov et al., 2015; Guzman et al., 1968; Huttly et al., 1990; Pradhan and Rawlings, 2002; Pronyk et al., 2012; Schlesinger et al., 1983; Stanton et al., 1988). Most of these studies were not randomized studies. One randomized trial from Bangladesh found no evidence of sanitation's impact on percent height-for-age, percent weight-for-age, or percent weight-for-age. Results from other non-randomized experimental studies are listed in the supplemental tables (S11–S13).

4.6.2. GRADE for intervention studies

Overall, the body of evidence examining the impact of sanitation interventions on stunting and underweight was of 'moderate' quality. Both sets of studies showcased minimal inconsistency, precise and protective pooled estimates of effect, and no risk of publication bias. However, the stunting studies had an average IQAT score of 6 and the underweight studies an average score of 5.2, indicating moderate risk of bias. Wasting intervention studies were given a GRADE of 'very low' quality of evidence since the average IQAT score was 3.5, suggesting high risk of bias, and the funnel plot showcased risk of publication bias.

4.6.3. Sanitation ladder and subgroup analyses

There were too few studies included for each of the nutrition outcomes to meaningfully assess movement up the sanitation ladder or sub-categories of age, season, region, follow-up time, water coverage, and soap coverage (Tables S25–S27).

5. Discussion

We reviewed the extensive evidence on the impact of sanitation on critical health outcomes, including diarrheal diseases, STH infections, trachoma, schistosomiasis, and nutritional status. We updated existing reviews, but included several new analyses to support policy decisions on sanitation and health. We relied on evidence from a large number of studies that included all population types and geographical regions to allow for generalizability of the results. While our conclusions are mostly derived from a separate analysis of intervention studies due to the greater rigor of this body of evidence, these studies still covered a broad range of populations and settings. We assessed the quality of the body of evidence based on intervention studies using GRADE. Overall, results suggest that access to sanitation facilities is protective, though pooled estimates were characterized by substantial heterogeneity. The quality of evi-
dence varied across the health outcomes from very low to high. Most studies followed observational designs and as a result, estimates for multiple outcomes were frequently taken from a small number of intervention studies and RCTs. Few studies were available that assessed movement up the sanitation ladder, limiting our ability to draw conclusions about the effectiveness of different levels of sanitation services. Too few studies reported on the use of sanitation (rather than access to sanitation) to draw any meaningful conclusions. Subgroup analysis revealed some evidence of impact among specific sub-populations, in particular age ranges, which support our understanding of the biology and transmission of STH, trachoma, and diarrheal diseases.

Evidence from observational and intervention studies suggest an association between sanitation on diarrheal disease consistent with findings from Wolf et al. (2014) that compared improved sanitation (but not sewered sanitation) versus unimproved sanitation (RR 0.84 95% CI 0.77, 0.91) (Wolf et al., 2014). This association held when the pooled analysis was restricted to intervention studies, though findings from studies in schools and communities varied and the quality of evidence was low. Heterogeneity may be a result of several factors that we were not able to capture in this review, including the key etiologic pathogens in circulation, and background rates of infection (Eisenberg et al., 2007). The role of handwashing following fecal contact as a complement to sanitation is critical, whereby sanitation may actually pose an individual risk of infection if toilets are not well maintained (Greene et al., 2012). However, few studies report on the complementarity of toilet maintenance or handwashing in the context of sanitation improvements. Water is an essential aspect of hand hygiene, and we found some evidence of a dose-response of water coverage in our analysis, whereby studies in areas with higher water coverage levels found a greater association between sanitation and diarrhea than in lower water coverage areas.

There was evidence for an association between sanitation and STH, but not when the analysis was restricted to intervention studies alone. We can conclude some evidence for an effect of sanitation on *A. lumbricoides*, as the measure of association was similar between all studies and intervention studies only, though the confidence intervals were much wider for intervention studies given the low number of studies available. We would expect the results to be consistent with *T. trichiura* given the similar biology and mechanisms of infection, but intervention effects were null for *T. trichiura*. There is a WHO target of achieving 75% school-based coverage of deworming medication in endemic areas; however, these drugs are more effective in clearing *A. lumbricoides* infections than other STHs (Steinmann et al., 2011). This may explain the stronger association of sanitation with *A. lumbricoides* as compared to the other STHs, as sanitation is unlikely to reduce infection over a short period of time, only reduce reinfection. In addition, infection with *A. lumbricoides* generally peaks among school-age population, so sub-groupings that found associations between sanitation and individual STH for that age group, as well as significant associations among school-based studies is consistent with the biology of disease. Sanitation may have a greater impact among those with higher rates of infection. Indeed, we did not find an association between school sanitation and hookworm. The effect of sanitation interventions on hookworm was much lower than the association between sanitation and hookworm derived from cross-sectional and observational studies. Given the transmission pathway for hookworm (through exposed skin in the feet and hands), shoe wearing may be a more important determinant of infection (Strunz et al., 2014). The previous review from Strunz et al. (2014) found stronger overall associations between sanitation and *A. lumbricoides* (OR 0.64, 95% CI 0.44, 0.88) and *T. trichiura* (OR 0.61, 95% CI 0.50, 0.74), but consistent null associations with hookworm (OR 0.80, 95% CI 0.61, 1.06).

We found strong evidence of an association between sanitation and active trachoma — measures of TF/TF — both through observational associations and intervention studies. The pooled estimates were larger than those found in the previous review by Stocks et al. (2014). However, the pooled estimate of the intervention studies was heterogeneous, with the study by Khandekar et al. (2006) reporting a very large effect, potentially overestimating the estimate of the pooled effect. Given the high rates of infection in younger age groups, most studies focus on pre- and school-age children, and associations were found in both groups. Too few studies assessed the sanitation ladder to provide guidance on association between specific sanitation technologies and trachoma. Few intervention studies assessed the impact on *C. trachomatis*, so the association with sanitation is less conclusive.

We found evidence consistent with the review conducted by Grimes et al. (2014) that sanitation is associated with lower odds of schistosomiasis, but the lack of intervention studies limits the quality of evidence and confidence in this association. Given the life cycles of the schistosomes, we would anticipate that access to improved water, predominantly used for bathing, might also play a considerable role in transmission (Grimes et al., 2015). Our data revealed associations with sanitation in studies from areas with both high and low improved water supply.

Our nutrition findings were similar to previous findings (Dangour et al., 2013) which reported no effect of sanitation on weight-for-age z-score (MD 0.05; 95% CI −0.15 to 0.25) or on weight-for-height z-score (MD 0.04; 95% CI −0.04 to 0.10), but a borderline effect on height-for-age z-score (MD 0.08; 95% CI 0.00 to 0.16). We found several rigorous randomized trials assessing these associations that were not included in the previous review, but the results from these trials were mixed. Because we followed the inclusion criteria of the previous review and only included intervention studies, too few studies were included to provide much guidance on subgroup analyses and movement up the sanitation ladder. Future updates should broaden their inclusion criteria to include observational studies.

As a complement to previous reviews, we assessed the impact of movement up the JMP sanitation ladder (WHO/UNICEF, 2016). There was limited evidence to suggest the benefit of improved sanitation over unimproved sanitation. Similarly, we did not find sufficient evidence to suggest the benefits of private versus shared sanitation beyond what has been reported in the previous meta-analysis by Heijnen et al. (2014). There was not enough consistency between intervention approaches to determine if any one particular approach to sanitation improvements led to greater improvements in health.

The results assessing the subset of studies that reported on latrine coverage and latrine use suggest that in order to observe sweeping health benefits, sanitation coverage levels probably have to be higher than what we observed in these studies. Recent studies suggesting herd protection among those living in communities with high sanitation coverage on diarrhea and other outcomes (Aldermon et al., 2003; Andres et al., 2014; Barreto et al., 2007; Fuller et al., 2016; Geruso and Spears, 2015; Spears et al., 2013). We point out that out coverage and use sub-analysis emphasized intervention studies that also recorded coverage and/or use, and that this population is a subset of the overall sanitation literature on health.

Most of the trials that we captured were studies of programmatic effectiveness, and coverage and use of sanitation, where measured, was low. Effectiveness trials assess the ability of a specific intervention under real-world conditions, and real-world conditions in sanitation interventions often means sub-optimal sanitation adherence. Evaluating interventions with low uptake or fidelity to sanitation tells us little about the importance of sanitation in preventing disease or adverse sequelae. There are very few...
efficacy trials in the sanitation literature (e.g., trials that assess the impact of sanitation under ideal adherence). Evidence on the health benefit of sanitation may be more compelling once evaluations are measuring the impact of successful efficacy interventions.

Sanitation serves as a primary barrier to mitigating both individual and communal fecal exposure, so assessments of individual exposure–disease relationships may be less meaningful than those assessing community-level coverage and use. However, our subgroup analysis of sanitation coverage and use was limited by the small number of intervention studies that also reported on coverage and use. Data on the role of increased community-level coverage and use for sanitation and community health outcomes may provide more biologically and policy-relevant data. To limit bias and confounding, this type of evidence requires experimental cluster-randomized designs, and very few of these studies are available.

There were several limitations to our review. While we reviewed thousands of articles assessing the impact of sanitation on myriad health outcomes, the body of evidence was by no means conclusive. Statistical heterogeneity in the meta-analyses was high throughout which may be a result of the substantial underlying heterogeneity in disease burden, existing water, sanitation, and hygiene coverage and other environmental conditions across the different studies, but also the different type of sanitation interventions being assessed. There was a lack of reporting for some of the variables that we used in the subgroup analyses (e.g., length of follow-up), and this could account for some of the unexplained heterogeneity, although often heterogeneity persisted even among the subgroup analyses of our available data. Sanitation, and in some cases health outcomes, were often measured differently across studies, making comparability, especially within our subgroup analysis challenging. We also could not always distinguish between the impacts of sanitation alone vs. sanitation combined with other components. The quality of evidence, as indicated by the GRADE criteria varied from moderate to very low for the intervention studies. It would have been ideal to assess bias among all observational studies, however, due to the large number of studies we focused the bias analyses only on experimental studies; It is likely that much of the observational study base would have scored poorly in such a bias assessment due to the high possibilities of confounding. Many of the trials that we found were program effectiveness studies, whereas efficacy studies (achieving high sanitation coverage and use) may be needed to demonstrate health impacts. Many of the trials also had only limited details on the implementation of their interventions, and future research should be more descriptive in their intervention details, including the persistence of the implementers in terms of time and number of visits, and the cost-effectiveness of the interventions. Of interest is how different changes in coverage and use (e.g., from 20% to 40% versus 40% to 80%) impact health effects and the relevant community coverage and use thresholds that would substantially reduce disease burden. Another of our group’s reviews specifically addresses how different types of sanitation interventions impact coverage and use (Carn et al., 2016), but there is further need to identify how coverage and use impact health.

While we did find protective associations with the STH and schistosomiasis outcomes and sanitation, these associations were not supported by the intervention studies (or the lack of intervention studies for schistosomiasis), limiting our confidence in associations. Findings of this review were consistent with the previous reviews we updated. The impact of sanitation on health is likely dependent on several factors, and some sanitation interventions may not do enough to mitigate possible routes of transmission. Background rates of disease, other water and hygiene characteristics of the population, the type of the sanitation intervention, and intervention fidelity, specifically use of the facilities and duration of follow-up are all likely to determine the role of sanitation in mitigating health burden.

Funding statement

The World Health Organization provided funding for this review. WHO funding was made possible through contributions from the UK Department for International Development and the Bill & Melinda Gates Foundation.

Competing interests

SB & KM are employees of the WHO. The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions or policies of the institutions with which they are affiliated. MCF consults for the WHO on various projects related to water, sanitation, and hygiene, and infectious disease. JVG consults for the WHO on a project related to water, sanitation, and hygiene, and trachoma infection. MCF is a member of the STH Advisory Board, which advises the WHO on STH policy. Emory University is a member of the WHO University Collaborating Counsel on Trachoma, which contributes to the WHO’s work plan on trachoma control.

Acknowledgements

We would like to acknowledge Rachel Stelmach, Henrietta Lewis, and Divya Narayan for their work on the title, abstract and full text reviews. We would like to acknowledge Heather Amato and Patrick Brooks for their assistance on data extractions, and Heather Amato for her assistance on the risk of bias assessments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijheh.2017.05.007.

References


