Background
Ever since John Snow's intervention on the Broad St pump, the effect of water quality, hygiene and sanitation in preventing diarrhea deaths has always been debated. The evidence identified in previous reviews is of variable quality, and mostly relates to morbidity rather than mortality.

Methods
We drew on three systematic reviews, two of them for the Cochrane Collaboration, focussed on the effect of handwashing with soap on diarrhoea, of water quality improvement and of excreta disposal, respectively. The estimated effect on diarrhoea mortality was determined by applying the rules adopted for this supplement, where appropriate.

Results
The striking effect of handwashing with soap is consistent across various study designs and pathogens, though it depends on access to water. The effect of water treatment appears similarly large, but is not found in few blinded studies, suggesting that it may be partly due to the placebo effect. There is very little rigorous evidence for the health benefit of sanitation; four intervention studies were eventually identified, though they were all quasi-randomized, had morbidity as the outcome, and were in Chinese.

Conclusion
We propose diarrhoea risk reductions of 48, 17 and 36%, associated respectively, with handwashing with soap, improved water quality and excreta disposal as the estimates of effect for the LiST model. Most of the evidence is of poor quality. More trials are needed, but the evidence is nonetheless strong enough to support the provision of water supply, sanitation and hygiene for all.

Keywords
Water, sanitation, hygiene, diarrhoea, mortality
water in diarrhoea transmission and prevention has been hotly debated. More recently, awareness has also grown about the importance of excreta disposal in preventing diarrhoeal disease, culminating in the recent poll of readers of the British Medical Journal in which sanitation was voted the greatest advance in public health in the last century.²

There is strong temptation to conduct evaluations of the health impact of water supply, sanitation and hygiene interventions, but the challenges also are many. Often it is difficult or impossible to randomize or to blind the intervention. In practice, most studies do not identify specific aetiologies and so deal with an outcome (diarrhoea) which is caused by various pathogens, transmitted by various routes and associated with various potential confounding factors. The vulnerability of such studies to confounding is compounded by the use of observational study designs and the low relative risks (RRs) involved, which are typically less than two. In developing countries, most episodes of diarrhoea morbidity—even much of the life-threatening morbidity—are not reported to the health system, so that active surveillance involving home visits is used to detect them, often with excessive recall periods. In the circumstances then, it is not surprising that the first methodological review of this literature³ located some 50 instances then, it is not surprising that the first methodological review of this literature³ located some 50 epidemiological studies, but found serious flaws in every one.

A series of literature reviews conducted by Esrey and others⁴–⁷ established a consensus view on the impacts on health of improved water quality, water quantity and sanitation, which was summarized in the relevant chapter of Disease Control in Developing Countries.⁸ A more recent review⁹ gave prominence to a number of studies of household-based water treatment, and arrived at a greater estimate of the impact of water quality than previous reviews. However, the confidence intervals (CIs) for this and the other such estimates were very wide, so wide as to show that the new figures were not significantly different from the corresponding previous estimates.

In this article, drawing on three systematic reviews of the literature, we present the evidence for an impact on diarrhoea mortality from improvements in hygiene (specifically, handwashing with soap),¹⁰ drinking-water quality,¹¹ and excreta disposal. The reviews of effectiveness of the interventions are shaped in large part by the needs of the LiST model. In that model, increases in coverage of an intervention result in a reduction of deaths due to one or more causes or in the reduction of a risk factor. Therefore, the reviews and the grade process used were designed to develop estimates of the effect of an intervention in reducing either a risk factor or deaths due to a specific cause. For more details of the review methods, the adapted grade approach or the LiST model, see other articles in this supplement.

**Methods**

**Handwashing with soap**

The original review¹⁰ aimed to identify all studies published in English up to the end of 2002 relating handwashing to the risk of infectious intestinal or diarrhoeal diseases in the community. Medline, CAB Abstracts, Embase, Web of Science and the Cochrane Library were systematically searched for papers related to handwashing, use of soap, as well as disease terms such as diarrhoea, typhoid, enteric, cholera, shigellosis, dysentery, and mortality. Searches were also undertaken by hand with reference lists from these papers, the authors' own collections and review articles. No limitations were placed on date or geographical location. The search was updated in 2008.

Studies were retained for the meta-analysis if they provided point estimates and 95% CIs (or the means to calculate them) of the risk of not washing hands. Intervention trials not solely concerned with handwashing were excluded. Where both crude and adjusted odds ratios were presented, adjusted values were used. The risk values for studies with several measures of handwashing practice were combined by averaging, if they concerned the same sample group. If they concerned different groups, they were treated as if they were separate studies. Similarly, studies with two different outcome measures were entered into the meta-analysis as if they were separate studies.

**Water quality**

Following the Cochrane peer-reviewed protocol,¹² we searched the specialized register of the Cochrane Infectious Diseases Group, CENTRAL, Medline, Embase and LILACS for all randomized and quasi-randomized controlled trials of interventions to improve water quality for the prevention of diarrhoeal disease, regardless of language, publication status, or date, up to December 2005. Interventions included any measure to improve the microbial quality of drinking water, unless undertaken in response to epidemic diarrhoea. The primary outcome was diarrhoea in adults or children. We hand searched conference proceedings, contacted researchers and organizations working in the specialty, and checked the references of identified studies. Two reviewers independently examined the electronic records for potentially eligible studies and the full text of potentially eligible reports. Disagreements were resolved by a third reviewer.

Measures of effect reported were risk ratios, rate ratios, odds ratios and longitudinal prevalence ratios (number of days or weeks with diarrhoea divided by number of days or weeks under observation in a person). The results are presented separately by study type and also by type of intervention, whether source- or household-based. A random effects inverse
variance method was used on the log scale to calculate pooled estimates.

**Excreta disposal**
Again we followed the Cochrane approved protocol\(^1\) to search for interventions to dispose of human excreta so as to reduce direct or indirect human contact. This includes any steps to remove or contain faeces, such as simple pit latrines, bucket latrines, hanging toilets and composting toilets, and should be contrasted with open defecation. Diarrhoea was again the outcome measure, whether or not microbiologically confirmed. We defined diarrhoea and an episode in accordance with the case definitions used in each trial. We excluded trials that had no clinical outcomes; for example, trials that only reported on microbiological pathogens in the stool.

We searched the following databases up to April 2007: Cochrane Infectious Disease Group Specialized Register; CENTRAL; MEDLINE; EMBASE; LILACS, and also Chinese-language databases available under the Wan Fang portal using the Chinese equivalents of our search terms where appropriate. We searched the metaRegister of Controlled Trials (mRCT) using ‘diarrhoea’, ‘diarrhoea’ and ‘sanitation or latrine or toilet or privy or disposal’ as search terms, and also a number of relevant conference proceedings. Other researchers and relevant international agencies were also contacted in the search for unpublished and ongoing trials. The reference lists of studies identified as above were also scanned for any further relevant studies.

**General**
All studies which met final inclusion and exclusion criteria were double data abstracted into a standardized form for each outcome of interest.\(^4\) We abstracted key variables with regard to the study identifiers and context, study design and limitations, intervention specifics, and outcome effects. Each study was assessed and graded according to the adaptation by the Child Health Epidemiology Reference Group (CHERG) of the GRADE technique.\(^5\) Studies received an initial score of high if a randomized or cluster randomized trial. The grade was decreased one grade for each study design limitation. In addition, studies reporting an intent-to-treat analysis or with statistically significant strong levels of association (>80% reduction) receive 1–2 grade increases. Any study with a final grade of very low was excluded on the basis of inadequate study quality. The process is summarized in Tables 1, 2 and 3. To save space, we have not listed all of the studies in all three reviews with the bibliography of this article; instead, a detailed listing is provided in three corresponding Supplementary Tables 1, 2 and 3 available at *LJE* online. The numbers in Tables 1, 2 and 3 refer to these supplementary tables.

For the outcome of interest, namely the reduction of diarrhoea mortality, we applied the CHERG Rules for Evidence Review\(^4\) to the collective diarrhoea morbidity and mortality outcomes to generate a final estimate of effect.

**Results**

**Handwashing with soap**
The search identified a total of 38 studies (Figure 1) but 21 were not suitable for data extraction, either because they did not specify whether soap was used, or did not present data permitting a calculation of effect. Of the 17 remaining studies, 7 were intervention trials and 10 were observational (Table 1). All of the interventions or exposure measures related to handwashing before eating or food handling, or after defecation or handling of child stools, or a combination of these.

Only one mortality study was found.\(^6\) The number of events was not stated, but the wide confidence intervals (reduction of +62% to −43%) suggest there were very few, and the study was observational, with weaknesses in the outcome definition (deaths from other infectious diseases besides diarrhoea were included) and ascertainment of compliance.

We therefore turned to morbidity studies. None of the intervention studies gave adequate compliance data, so that the effect reported (a reduction of 47%) is that of handwashing promotion rather than of handwashing itself; i.e. the effect of the intervention, not the individual’s response to it. However, this was slightly greater than the pooled effect (a reduction of 43%) of all studies in the review (Figure 2a and b). In other words, the observational studies, which did report the effect of individuals’ behaviour, did not find a greater effect. The original review found that the effect was remarkably consistent across studies of higher methodological quality, and studies with severe forms of diarrhoea as an outcome.

In each of these categories, the pooled estimate of effect is of a reduction within the range of 42–48% (Table 1). The combined effect of the more severe diarrhoeas was a 48% reduction. Two studies of laboratory confirmed shigellosis,\(^7,8\) a more severe and a ‘harder’ outcome than self-reported diarrhoea, had a pooled effect of reduction by 59%. On the other hand, three studies were conducted in settings where water use was known to be constrained; (i) a low-income area of Lima, Peru, where vendors sell water expensively from tanker trucks, (ii) a refugee camp in Malawi and (iii) a setting in Burundi where median water usage was only 51 per capita per day. The reductions in risk were 11, 26 and 41%, respectively, all of them less than the combined effect of 43% found in the review as a whole.

Subsequent to our initial review,\(^9\) a Cochrane review of the effect of handwashing on diarrhoea was published.\(^9\) Most of the studies included were in institutional settings, but five were conducted in the community. In two of those, the intervention
Table 1  Quality assessment of trials of handwashing with soap for the prevention of diarrhoea

<table>
<thead>
<tr>
<th>No of studies (supplementary table ref.)</th>
<th>Design</th>
<th>Limitations</th>
<th>Consistency (based on the heterogeneity of the meta)</th>
<th>Generalizability to population of interest</th>
<th>Generalizability of intervention</th>
<th>No of events</th>
<th>Intervention</th>
<th>Control</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome 1: Diarrhoea mortality; Quality: very low</strong></td>
<td>Observational (case control)</td>
<td>Not randomized, no placebo, unreliable exposure measure (−2)</td>
<td>Not applicable</td>
<td>Only one study (−0.5)</td>
<td>Not given</td>
<td>Not given</td>
<td>15</td>
<td>15</td>
<td>0.97 (0.38–1.43) *</td>
</tr>
<tr>
<td><strong>Outcome 2: Diarrhoea morbidity; Quality: very low</strong></td>
<td>RCT/quasi-RCT</td>
<td>Inadequate randomization, no placebo, compliance not measured, no baseline incidence (−1)</td>
<td>Tests for heterogeneity gave significant result (−0.5)</td>
<td>1 study in USA, 1 Australia; 3 studies are all ages (−0.5)</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>739^b</td>
<td>1157^b</td>
<td>0.53 (0.37–0.76) ^c</td>
<td></td>
</tr>
<tr>
<td><strong>Outcome 3: Severe outcomes (hospitalized enteric infection cases, cholera, shigellosis, typhoid and deaths); Quality: very low</strong></td>
<td>RCT/quasi-RCT/observational</td>
<td>Inadequate randomization, no placebo, compliance not assessed, no baseline incidence (−1)</td>
<td>Tests for heterogeneity gave significant result (−0.5)</td>
<td>Mostly Asia; most studies are all ages (−0.5)</td>
<td>Severe morbidity and cause-specific mortality (−0.5)</td>
<td>492^b</td>
<td>798^b</td>
<td>0.52 (0.34–0.65) ^c</td>
<td></td>
</tr>
<tr>
<td><strong>Outcome 4: Shigellosis; Quality: low</strong></td>
<td>RCT/quasi-RCT</td>
<td>Inadequate randomization, no placebo, compliance not measured, no baseline incidence (−1)</td>
<td>Heterogeneity not significant. Random effects models used for pooled estimates. Both fixed and random effect models gave the same pooled estimate</td>
<td>Both studies in Asia, both all ages (−0.5)</td>
<td>Severe morbidity</td>
<td>17</td>
<td>51</td>
<td>0.41 (0.27–0.62) ^c</td>
<td></td>
</tr>
</tbody>
</table>

Source of systematic review and measures of effect: Ref.10

*Effect comes from unadjusted odds ratio of 0.97 reported in Hoque et al.16

^bNumber of events not available for all studies.

^cRandom effect meta-analysis.

^dIncludes trials with baselines and concurrent control groups and observational studies with adequate control for confounding.
Table 2  Quality assessment of trials of water quality improvement for the prevention of diarrhoea

<table>
<thead>
<tr>
<th>No of studies (supplementary table ref.)</th>
<th>Design</th>
<th>Limitations</th>
<th>Consistency (based on the heterogeneity of the meta)</th>
<th>Generalizability to population of interest</th>
<th>Generalizability of intervention</th>
<th>No of events</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome 1: Diarrhoeal mortality for all ages; Quality: very low</strong>&lt;sup&gt;26,27&lt;/sup&gt;</td>
<td>RCT and quasi-RCT</td>
<td>No placebo, not blinded, one randomized, self-reported and self-defined diarrhoea (−1.5)</td>
<td>NA</td>
<td>Both Africa, 1 is all ages, 1 under 5s (−0.5)</td>
<td>1 all-cause mortality, 1 diarrhoeal mortality</td>
<td>34</td>
<td>55</td>
</tr>
<tr>
<td><strong>Outcome 2: Diarrhoeal morbidity for all ages; Quality: very low</strong>&lt;sup&gt;38 (30 studies, 5 with multiple arms; arms counted as additional trials)&lt;sup&gt;1−30&lt;/sup&gt;</td>
<td>RCT and quasi-RCT</td>
<td>Only three blinded and placebo, few adequately randomized, most self-reported diarrhoea with excessive recall period. (−1.5)</td>
<td>Test for heterogeneity gave significant result (−0.5)</td>
<td>Mostly Africa, Latin America and Asia, all ages (−0.5)</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>41 485</td>
<td>41,5,12 Double blind studies</td>
</tr>
<tr>
<td>27 (19 studies, 5 with multiple arms)&lt;sup&gt;1−19&lt;/sup&gt;</td>
<td>RCT</td>
<td>Only three blinded and with placebo, few adequately randomized (−1)</td>
<td>Heterogeneity not significant (P = 0.30)</td>
<td>1 Lat Am, 1 Africa, 1 USA, all ages (−0.5)</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>1912</td>
<td>1987</td>
</tr>
<tr>
<td>11&lt;sup&gt;20−30&lt;/sup&gt;</td>
<td>Quasi-RCT</td>
<td>Inadequately randomized, no placebo, mostly self-reported with excessive recall period. (−2)</td>
<td>Significant heterogeneity (−0.5)</td>
<td>Mostly Asia and Africa, all ages (−0.5)</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>18 334&lt;sup&gt;b&lt;/sup&gt;</td>
<td>19 125&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>6&lt;sup&gt;20,21,23,24,27,30&lt;/sup&gt;</td>
<td>Quasi-RCT source-based interventions</td>
<td>Inadequately randomized, no placebo, mostly self-reported, excessive recall period (−2)</td>
<td>Significant heterogeneity (−0.5)</td>
<td>4 Asia, 2 Africa, all ages (−0.5)</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>18 164</td>
<td>18 273</td>
</tr>
<tr>
<td>32&lt;sup&gt;1−19,22,25,26,28,29&lt;/sup&gt;</td>
<td>Quasi-RCT household-based interventions</td>
<td>Only three blinded and placebo, few adequately randomized, mostly self-reported diarrhoea with excessive recall period. (−1.5)</td>
<td>Significant heterogeneity (−0.5)</td>
<td>Mostly Africa and Latin America, all ages (−0.5)</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>23 321&lt;sup&gt;b&lt;/sup&gt;</td>
<td>45 500&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Outcome 3: Diarrhoeal morbidity for children &lt;5 years; Quality: very low</strong>&lt;sup&gt;29&lt;/sup&gt;&lt;sup&gt;1−5,11−13,20−22,24,26,27&lt;/sup&gt;</td>
<td>RCTs and quasi-RCTs</td>
<td>Only three blinded and placebo, few adequately randomized, most self-reported diarrhoea with excessive recall period (−1.5)</td>
<td>Significant heterogeneity (−0.5)</td>
<td>Mostly Latin America and Africa, under 5s</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>17 687&lt;sup&gt;b&lt;/sup&gt;</td>
<td>39 165&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>4&lt;sup&gt;20,23,24,27&lt;/sup&gt;</td>
<td>Quasi-RCT source-based interventions</td>
<td>Inadequately randomized, no placebo, mostly self-reported, excessive recall period (−2)</td>
<td>Test for heterogeneity gave significant result (−0.5)</td>
<td>2 Africa, 2 Asia, under 5s</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>1603</td>
<td>1924</td>
</tr>
<tr>
<td>25&lt;sup&gt;1−4,8,11−13,15−19,26,27&lt;/sup&gt;</td>
<td>RCT and quasi-RCT household-based interventions</td>
<td>Only three blinded and placebo, few adequately randomized, most self-reported diarrhoea with excessive recall period. (−1.5)</td>
<td>Significant heterogeneity (−0.5)</td>
<td>Mostly Latin America and Africa, under 5s</td>
<td>Moderate and severe morbidity (−0.5)</td>
<td>16 084&lt;sup&gt;b&lt;/sup&gt;</td>
<td>37 241&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Source: Clasen T, 2006, PhD Dissertation, University of London.<sup>11</sup>

<sup>a</sup>We did not calculate pooled estimates for the studies on water quality reporting mortality.

<sup>b</sup>Number of events not available for one or more studies.

<sup>c</sup>Random effects inverse variance method on the log scale was used to calculate pooled estimates. Heterogeneity was examined using the χ² test with a 10% level of statistical significance and the I² test for consistency.

<sup>1</sup>...<sup>11</sup>...
Water quality interventions

Two studies with a mortality outcome were found and included, following the Cochrane protocol (see Figure 3 and Table 2). One of these had ascertained diarrhoea mortality, but the intervention included health education about oral rehydration therapy for diarrhoea, which alone could explain the 85% reduction in diarrhoea mortality that was observed. Moreover, only two villages had been randomized. The other trial, whose authors admit that it was not designed to detect a mortality outcome, found a significant reduction (RR of death = 0.58, P < 0.036) only for all causes and all ages by pooling two intervention arms (total 59 events). Pooling the studies would be questionable, as the interventions were very different.

The serious limitations of the mortality studies led us to consider the morbidity studies. All took self-reported diarrhoea as the outcome; the Cochrane protocol had excluded trials that had no clinical outcomes, such as trials that report only on microbiological pathogens in the stool. Similar results were obtained for diarrhoea in all age groups and with the outcome limited to children aged <5 years. The type of study seemed to have little bearing on the measure of effect, which ranged from 46% in randomized controlled trials (RCTs) to 38% of quasi-RCTs. The review protocol approved by the Cochrane Collaboration had limited its scope to trials, so that no observational studies were included. Interventions based at the water source had less effect (27% for all ages, 15% for young children) than household-based ones, with an effect of 43–44%, depending on the age range. Two studies implemented in settings where the water before treatment had ≤10 faecal coliforms/100 ml (a concentration classified by WHO as
Figure 2 Forest plots of (a) all studies in handwashing review, and (b) intervention studies only. Numbers on y-axis are references to studies in Supplementary Table 1. The diamond represents the combined relative risk and 95% CI from random effects model.\textsuperscript{10}
‘low risk’) had a combined effect of reducing risk by 61% (95% CI 30–67%) with insignificant heterogeneity. The greatest difference of all was between the 31 unblinded and the 4 blinded trials, 25–28 with the latter giving a reduction of only 7% in diarrhoea risk, which was not statistically significant (Figure 4).

Excreta disposal
Our initial attempt at this review nearly foundered. The initial search13 produced seven quasi-randomized intervention studies with diarrhoea morbidity as the outcome measure, and one with diarrhoea mortality (Figure 5, Table 4). The mortality study has already been mentioned;23 the intervention included water supply and education about oral rehydration therapy, which could explain the observed reduction in mortality. All eight studies involved an intervention to improve water supply as well as excreta disposal, so that in none of them is it possible to assess the impact of sanitation as an intervention per se.

A renewed effort to locate suitable studies produced a further four,29–32 in which the intervention involved excreta disposal alone. All four were conducted in China (see the following URL http://maps.google.com/maps/ms?ie=UTF8&hl=en&msa=0&msid=109120590372695346058.000465c8157b73053ba7f&z=6 for a map of the study sites), and published in Chinese (Table 3). The reductions in diarrhoea morbidity in the four studies were 63, 51, 20 and 8%, respectively. In all but the last, the confidence interval did not include zero reduction. We did not calculate pooled estimates for the sanitation trials, because most studies randomized a very small number of villages, which makes calculation of confidence intervals impossible. Also, the interventions tested were very different from one another, which also makes pooling questionable.

These four studies are hardly ideal; they are quasi-randomized, not full RCTs; the control groups did not lack sanitation altogether, but mainly used some sort of pit latrine; and third, the published studies are not available for scrutiny except by those able to read Chinese. We therefore widened our scope further to include before/after studies. The unit of intervention is effectively the community, neighbourhood or village, but only one such study considered enough such units for statistical tests to be applied at this level. This recent study33 is, strictly speaking, a before/after design, but is in many ways akin to a

Figure 3 Synthesis of study identification in review of the effect of water quality interventions on diarrhoea mortality and morbidity11

Figure 4 Forest plots of (a) all 35 studies in water quality review, and (b) the four blinded studies only35
quasi-randomized trial in 24 neighbourhoods of the City of Salvador, Brazil. It found a city-wide reduction of 21% (95% CI 19–26%), and 43% (CI 39–46%) in the high risk areas.

Discussion

Handwashing with soap

Once the nature of the studies has been consistently defined (in the present case, community-based studies with an intervention focussed on handwashing and involving soap), the studies in the literature are remarkably consistent, showing a reduction in diarrhoea by 42–48%. We took 48%, the reduction found for the more severe diarrhoeas, as the figure to propose for LiST.

The problem encountered with water quality interventions, that the blinded studies do not support the positive picture drawn by the others, raises the question of whether this problem applies to handwashing with soap. After all, one cannot persuade people to wash their hands without their knowledge! The differing results of the blinded and unblinded studies suggest that much of the apparent impact of water treatment is attributable to bias, but by the same token that possibility cannot be ignored in the case of handwashing.

Blinding can be applied to two groups; the subjects, and those who assess their status. It can refer to each subject’s exposure allocation, or to the nature of the exposure or the disease outcome. Indeed, the Cochrane review of handwashing19 noted that one

![Figure 5](http://ije.oxfordjournals.org/)

**Figure 5** Synthesis of study identification in review of the effect of excreta disposal on diarrhoea morbidity
study had used a placebo intervention to conceal from the subjects and assessors which was the intervention group and which the control, and had found a smaller effect than the other community-based trials. An alternative explanation is that this study did not focus on handwashing or provide soap. All of the community-based trials which did not focus on handwashing or provide soap found smaller reductions in diarrhoea than either of those which did those things. Of the less focussed group, the partially blinded study found a reduction of 6% (95% CI: −44–15%) while the other, open trial of this kind found 25% (95% CI 15–34%) so the lack of blinding seems to explain at least part of the difference. Of course, such non-focussed trials were excluded from our earlier review, and none of the intervention studies focussed on handwashing (Figure 2a) used a placebo intervention. The consistency of effect between observational and intervention studies led us to judge that the effect is probably genuine, but more research is needed to clarify this point.

**Water quality interventions**

Enthusiasts for point-of-use household-based water treatment argue that the smaller effect of source-based interventions is due to subsequent contamination of the water on its way to, or during storage in the household. Sceptics point to the lack of significant effect found in the blinded studies, attributing the difference to courtesy bias or placebo effect. They suggest that consumers are less conscious of treatment administered centrally, and so less likely to show these forms of bias in their self-reporting of diarrhoea.

These findings have given rise to a lively debate about the desirability of scaling-up household-based water treatment. The only certain conclusion is that the implications are uncertain, and it is unlikely that a Delphi process would arrive at an amicable consensus.

There are several reasons to believe that the bias may apply to water quality but which are less applicable to handwashing. First, there is biological plausibility; in many settings, a subject would have to ingest very large amounts of water in order to consume an infectious dose of a bacterial pathogen. Second, the anomaly that the reduction in diarrhoea seems to be independent of the quality of the ambient water before it is treated. Third, the observational studies of drinking-water quality (e.g. see ref. 38) do not show such large effects as the point-of-use intervention trials. And fourth, most of these trials were funded by manufacturers of water treatment chemicals or equipment.

In these circumstances the data from the recent spate of trials—and hence, any systematic review based on them—do not offer a firm basis for judging the effect of water quality improvements. The options include adopting the effect of the four blinded trials (no significant reduction in any trial), adopting the pooled effect of the source-based intervention studies (three trials among children under five gave a 7% reduction), or keeping the pre-existing consensus view arrived at two decades ago in reviews of observational and source-based intervention studies, that water quality improvements can be expected to be associated with a reduction of some 17% in diarrhoea risk. This latter figure was proposed for LI$^*$ as associated with use of an improved water supply, which is available within a reasonable distance, requiring a return journey of 30 min or less.

One theme which emerged from the debate was that ‘harder’ outcomes such as care-seeking for diarrhoea would be more objective and less prone to bias than a relatively ‘softer’ outcome such as self-reported diarrhoea. One well-known water quality trial excluded from our review because the outcome included gastrointestinal symptoms and of hospitalizations was similar in both groups. In this context, one recent trial with both handwashing and water quality arms is of particular interest, as the authors documented the effect of each arm both in terms of self-reported diarrhoea and of care-seeking (see Table 5). The effect of each intervention is very similar, whichever the outcome. This suggests, either that any bias affecting the former outcome also affected the latter; or that the former was not susceptible to bias as the blinded studies suggest.

<table>
<thead>
<tr>
<th>Intervention group</th>
<th>Reduction in weekly self-reported diarrhoea prevalence (95% CI)</th>
<th>Reduction in care-seeking for diarrhoea in children &lt;5 years (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soap and hand washing promotion</td>
<td>45% (12–68%)</td>
<td>48% (15–71%)</td>
</tr>
<tr>
<td>Bleach water treatment</td>
<td>53% (22–75%)</td>
<td>54% (22–77%)</td>
</tr>
<tr>
<td>Flocculant-disinfectant water treatment</td>
<td>59% (29–82%)</td>
<td>61% (31–84%)</td>
</tr>
<tr>
<td>Flocculant-disinfectant plus hand washing with soap</td>
<td>50% (18–72%)</td>
<td>55% (23–77%)</td>
</tr>
</tbody>
</table>
Excreta disposal
It is in this area that the evidence is weakest. It is not surprising that trials are few, as provision of hundreds of latrines is expensive. Observational studies cannot be trusted, in view of evidence that people in latrine-owning households (a self-selected group) behave more hygienically than others, even in respects which have nothing to do with excreta disposal.41,42 This is not to say that excreta disposal has no effect on diarrhoea. Indeed, there is a striking consistency between the reductions found in various reviews of 36%,7 32%,9 20–51% (the four Chinese studies) and 22–43%.33 That being so, there is not enough evidence to justify a departure from the prevailing consensus, published nearly two decades ago7 and widely cited with approval since then, that sanitation reduces diarrhoea risk by about 36%. This then is the strength of effect proposed for the LiST model.

General
The lack of conclusive evidence of the effect on diarrhoea of improvements in water, hygiene and sanitation in developing countries is not an excuse for inaction.43 We know enough to do a lot of good. In spite of doubts about the detail, it is clear that such environmental interventions can have a substantial effect on diarrhoea morbidity, and the very few relevant studies44 confirm that they have a similar effect upon diarrhoea mortality. Moreover, water, hygiene and sanitation have other important benefits, including the emancipation of women from drudgery and the enhancement of human dignity, and even other health benefits such as the control of trachoma and of intestinal helminths.

It is not entirely surprising that the evidence is weak; studies of mortality are fraught with ethical and logistic problems; water supplies and sanitation are expensive interventions to trial; and the engineers who install them are not accustomed to trials as an important part of their professional culture. Moreover, it is particularly difficult to blind a trial of an intervention involving the provision of hardware or the promotion of behaviour change, and as we have seen, it appears that the lack of blinding can lead to substantial bias if the outcome is ‘soft’.

In the context of a literature of uneven quality such as this, these reviews can also teach us that there is no definitive systematic review, or objective set of rules for conducting one. The examples here show how the result of a systematic review depends upon a number of questions, such as the definition of the outcome (should only diarrhoea be accepted, and vomiting rejected?), of the intervention (is sewerage in Brazil different from pit latrines in Africa?), on the range of languages accepted (in several recent reviews,13,45 more than a third of useful studies were in Chinese) and on which are regarded as the gravest methodological deficiencies. Each of those questions has more than one reasonable answer. Judgement-free data are a myth.

Supplementary data
Supplementary data are available at IJE online.

Funding
The US Fund for UNICEF from the Bill & Melinda Gates Foundation (grant 43386, partially) to ‘Promote evidence-based decision making in designing maternal, neonatal and child health interventions in low- and middle-income countries’ and UNICEF, Unilever research and WaterAid.

Acknowledgements
Stimulating discussions with colleagues in the CHERG have helped to improve this article; their input is gratefully acknowledged. So are the major contributions of Tom Clasen as principal author of the reviews on water and on sanitation, the patient and constructive support of Therese Dooley and the administrative wizardry of Eileen Chappell.

Conflict of interest: None declared.

KEY MESSAGES
- Effect of handwashing with soap is most consistent at roughly 48% reduction in diarrhoea.
- Effect of water quality improvements found in RCTs seems to be affected by bias – not seen in blinded studies.
- Evidence for effect of sanitation is weakest – randomized trials are needed – but may be 36% reduction.
- Though evidence is weak compared with clinical RCTs, it is enough for action.
- Analysing such evidence needs more than algorithms – it requires judgement.
References


40 Luby SP, Agboatwalla M, Painter J *et al.* Combining drinking water treatment and hand washing for diarrhoea prevention, a cluster randomised controlled trial. *Trop Med Int Health* 2006; **11**: 479–89.