



# Outcome reporting bias in evaluations of public health interventions: evidence of impact and the potential role of a study register

Mark Pearson, Jaime Peters

► Additional materials are published online only. To view these files please visit the journal online (<http://jech.bmj.com/content/66/4.toc>).

Peninsula Technology Assessment Group (PenTAG), Peninsula Medical School, University of Exeter, Exeter, UK

## Correspondence to

Mark Pearson, Peninsula Technology Assessment Group (PenTAG), Peninsula Medical School, University of Exeter, Veysey Building, Salmon Pool Lane, Exeter EX2 4SG, UK; [mark.pearson@pms.ac.uk](mailto:mark.pearson@pms.ac.uk)

Accepted 7 April 2011  
Published Online First  
7 June 2011

## ABSTRACT

**Background** Systematic reviews of the effectiveness of interventions are increasingly used to inform recommendations for public health policy and practice, but outcome reporting bias is rarely assessed.

**Methods** Studies excluded at full-text stage screening for a systematic review of a public health intervention were assessed for evidence of study exclusion resulting from non-reporting of relevant outcomes. Studies included in the review were assessed for evidence of outcome reporting bias and the impact that this had on the evidence synthesised using a formal tool (Outcome Reporting Bias in Trials (ORBIT)).

**Results** None of the reports excluded at full-text stage were excluded because of non-reporting of relevant outcomes. Of the 26 included papers, six were identified as having evidence of missing or incompletely reported outcomes, with 64% of unreported or incompletely reported outcomes identified as leading to a high risk of bias according to the ORBIT tool. Where there was evidence of the effectiveness of interventions before an assessment of outcome reporting bias was undertaken, identifying possible instances of outcome reporting bias generally led to a reduction in the strength of evidence for the effectiveness of the interventions.

**Conclusion** The findings from this single evaluation provide empirical data to support the call for a prospective public health interventions study registry to aid the identification of unreported or incompletely reported outcomes. Critical appraisal tools can also be used to identify incompletely reported outcomes, but a tool such as ORBIT requires development to be suitable for public health intervention evaluations.

## INTRODUCTION

Systematic reviews of the effectiveness of interventions are increasingly used by decision-making bodies as a source of evidence to inform recommendations for public health policy and practice. While the rigour of studies included in public health systematic reviews is routinely assessed using a quality appraisal tool,<sup>1–3</sup> with the exception of the Cochrane risk of bias tool,<sup>1</sup> none assesses potential bias to a subsequent review from missing or incompletely reported results. Outcome reporting bias occurs when outcomes are selectively reported on the basis of their statistical significance (eg,  $p < 0.05$ ), the magnitude of their estimated effect or their perceived interest to the study authors or intended readers.<sup>2,3</sup> Selective outcome reporting is similar to publication bias, but rather than having a whole study missing from the literature, selective

outcome reporting describes the situation where certain outcomes within a study are missing.

The lack of appraisal of outcome reporting bias in the public health literature may be due to the wide range of study designs that are appropriate to assess complex public health interventions.<sup>4–6</sup> The resulting diversity in reporting practice<sup>7</sup> makes the assessment of outcome reporting bias considerably more difficult. However, there is considerable evidence that outcome reporting bias exists in randomised controlled trials (RCTs)<sup>8,9</sup> and impacts on systematic reviews of RCTs.<sup>10</sup> A potential solution is the prospective registration of RCTs, backed up by journal editorial policies, which make publication conditional on prospective registration having taken place. Recently, a call for a prospective public health interventions study register reflecting the range of study designs appropriate to the evaluation of complex public health interventions was made.<sup>11</sup> As with selective reporting in RCTs, this would help overcome the problem of reporting bias in public health studies. The aims of this paper were to demonstrate the use of a formal tool (Outcome Reporting Bias in Trials (ORBIT)<sup>10</sup>) to appraise studies included in a systematic review of public health interventions for the existence of outcome reporting bias and to assess the impact that this bias (where present) had on the evidence that was used to inform the development of guidance. The systematic review<sup>12</sup> included 26 papers, reporting 22 studies (10 RCTs, three cluster RCTs, four controlled before and after studies and five uncontrolled before and after studies).

## METHODS

We purposively selected a systematic review of a public health intervention in which one of us (MP) was the lead author<sup>12</sup> in order to assess the included studies for evidence of outcome reporting bias and the impact that this had on the evidence synthesised. First, we assessed all the papers excluded at full-text stage (at which screening for inclusion in the review occurred;  $n=154$ ) for evidence of study exclusion resulting from non-reporting of relevant outcomes. A further 5272 abstracts were not assessed in this way because they were clearly excludable as not relevant to the review question. Second, we critically re-read each of the 26 papers included in the systematic review to identify missing or incompletely reported outcomes. For example, in our critical re-reading, we attempted to reconcile the reporting in the methods and results sections to determine whether an outcome was likely to have been measured and

analysed but not reported. Once we had identified papers with missing or incompletely reported outcomes, we (MP and JP) discussed three papers with respect to the ORBIT tool and the classification of missing or incompletely reported outcomes as low, high or no risk of bias.<sup>10</sup> The remaining 10 papers were classified by one of the authors (JP) for instances of missing or incompletely reported outcomes using the ORBIT tool. Each classification was then discussed between the two authors (MP and JP) in order to assure consistency or to reach a consensus. Difficulties encountered in applying the ORBIT tool to the public health intervention studies meant that all 10 of these papers (in addition to the three initially discussed) required in-depth discussion in order to reach a consensus on the classification of outcome reporting bias.

We applied the ORBIT classification system<sup>10</sup> to assess (1) the level of reporting of the outcome data (partial or none), (2) the likely risk of bias to each review from the exclusion of these data (no, low or high risk) and (3) the impact that inclusion of these data may have had on the findings of the review. The risk of bias is informed by the assumed reasons for not reporting the full outcome. If the non-reporting is considered to be related to the results of an analysis (eg, a statistically non-significant effect is observed), this leads to an assumption of there being a high risk of bias by not including this piece of evidence in the review. The risk of bias is defined by Kirkham and colleagues<sup>10</sup> as 'arising from the lack of inclusion of non-significant results when a trial was excluded from a meta-analysis or not fully reported in a review because the data were unavailable'.

We tabulated our assessment by type of outcome (eg, injuries, installation of home safety equipment, knowledge and behaviour) alongside the results of included studies where no instances of missing or incompletely reported outcomes were identified (see online-only material file 1). This enabled us to assess the impact on review findings due to missing or incompletely reported outcomes. In particular, we wanted to know whether the findings from the initial review would change in the light of the assessment for selective reporting within studies. This was achieved by making a judgement on the weight and strength of evidence. Statistical methods could not be used (as in Kirkham

*et al*<sup>10</sup>) as these are based on the premise that a meta-analysis is undertaken, but a meta-analysis of the public health studies was not deemed appropriate for this review.

## RESULTS

Not one of the 154 reports excluded at full-text stage were excluded because of non-reporting of relevant outcomes. Of the 26 papers included in the original systematic review (providing 58 estimates of the effectiveness of different interventions), six were identified as having evidence of missing or incompletely reported outcomes, contributing to a total of 19 instances of missing or incompletely reported outcomes (table 1).

According to the ORBIT rating, 12 of the 19 instances (64%) of unreported or incomplete reporting of outcomes were identified to lead to a high risk of bias to the findings of the review. The most common classification for incomplete reporting or non-reporting of outcomes in the review was ORBIT rating E (37%): judgement indicates that outcome was measured and analysed but not reported due to its results. In 26% of the 19 instances of unreported or incompletely reported outcomes, results were only reported as 'not significant' (typically stating  $p>0.05$ ) (ORBIT rating A). Only three instances (16%) of missing or incompletely reported outcomes were assessed to be of low risk to the review because of their exclusion (all ORBIT rating F). However, four instances (18%) of missing or incompletely reported outcomes were assessed to not lead to bias in the review (all ORBIT rating B).

The impact of these missing data on the findings of the review is shown in table 2. More detailed information on the individual study results and assessment of outcome reporting bias can be found in online-only material file 1. Where evidence for an effect of interventions had been identified in the original review,<sup>12</sup> identifying possible instances of selective reporting with a high risk of bias due to its exclusion generally led to a reduction in the strength of evidence for the effectiveness of the interventions. Assessing for outcome reporting bias resulted in moderating the evidence of intervention effects only in two injury types: one was moderated from strong evidence of effect to some evidence of effect (home safety knowledge about prevention of scalds)

**Table 1** Classification of instances of no or incomplete reporting in systematic review of public health interventions to reduce unintentional injuries to children in the home

ORBIT classification	ORBIT assessment of risk of bias	Number (%) of instances of unreported or incomplete reporting of outcomes	Example
A	High	5 (26)	Posner <i>et al</i> <sup>13</sup> only report results for the intervention versus control group comparison for water safety as $p<0.33$ (it is likely they mean $p=0.33$ ). No further information is provided to allow calculation of the effect estimate beyond the p value.
B	None	4 (21)	Posner <i>et al</i> <sup>13</sup> only report results for the intervention versus control group comparison for poison safety score as $p<0.02$ . No further information is provided to allow calculation of the effect estimate beyond the p value.
C	Low	0 (0)	
D	High	0 (0)	
E	High	7 (37)	Paul <i>et al</i> <sup>14</sup> collected data on the installation of lockable cabinets for poisons, but the results are not reported separately for the intervention and control group. Judgement suggests that this is due to statistically non-significant results since all reported results are statistically significant.
F	Low	3 (16)	In Mallonee <i>et al</i> , <sup>15</sup> comparison of smoke alarm installation is only made pre- and postintervention in the intervention group, even though smoke alarm installation data for a control group were reported to have been collected. However, the focus of the paper was on pre- versus postintervention comparison in the intervention group and so lack of comparison with the control group is not believed to be due to results of such a comparison.
G	High	0 (0)	
H	Low	0 (0)	
I	None	0 (0)	
Total		19	

ORBIT, Outcome Reporting Bias in Trials.

**Table 2** Strength of evidence before and after an assessment of possible reporting bias

Change	Conclusion before assessing for reporting bias	Conclusion after assessing for reporting bias	Outcomes
Moderated	Strong evidence of effect	Some evidence of effect	Home safety knowledge (scalds)
Moderated	Some evidence of effect	Inconclusive	Home safety (parental self-efficacy)
Unchanged	Strong evidence of effect	Strong evidence of effect	Home safety knowledge (poisoning)
			Smoke alarm installation
Unchanged	Some evidence of effect	Some evidence of effect	Electric socket cover installation
Unchanged	Inconclusive	Inconclusive	Injuries (fire-related)
			Poison cupboard latch installation
			Home safety knowledge (fires)
			Home safety knowledge (falls)
			Home safety knowledge (drowning)
			Home safety knowledge and behaviour (general)

and one from some evidence of effect to inconclusive results (parental self-efficacy for home safety) (see table 2). Interventions for the six injury types that were inconclusive in the original review remained so after assessment using the ORBIT tool (see table 2). For home safety knowledge type (prevention of poisoning) and smoke alarm installation, strong evidence for the effectiveness of interventions remained after an assessment of selective reporting was made. The evidence for the effectiveness of interventions for electric socket cover installation remained after investigation of selective reporting.

**DISCUSSION**

This study demonstrates that statistical significance is the most common cause of reporting bias in a review of public health studies. Consequently, most instances of unreported or incomplete reported outcomes identified here were rated as leading to a high risk of bias in the review. The likely impact of the unreported or incompletely reported outcomes was mixed: it either had little impact or reduced the strength of evidence for an effect of the intervention. The latter is similar to the findings from Kirkham and colleagues<sup>10</sup> that effectiveness estimates had been overestimated.

Many limitations of this study relate to the heterogeneity of the studies included in the review, both in terms of their design and their reporting, which hindered (1) the identification of missing or incompletely reported outcomes, (2) using ORBIT to classify instances of missing or incompletely reported outcomes

as to their potential bias to the review, and (3) assessing the impact of such instances on the conclusions of the review. A substantial degree of judgement was required to address these issues.

Unlike Kirkham and colleagues,<sup>10</sup> we did not involve public health experts. However, we believe our general conclusion (that effectiveness is over-estimated when reporting bias is not considered) would not change had experts been involved. Furthermore, we only report the findings from an evaluation of within-study reporting bias in one review. However, the findings from this single evaluation provide empirical data to support the call for a prospective public health interventions study registry.<sup>11</sup> A prospective study registry would aid the identification of unreported or incompletely reported outcomes<sup>16</sup> and also promote the cost-effective use of research funds by reducing unnecessary duplication of studies.<sup>11</sup> The evidence required to inform public health policy that addresses the determinants of health inequalities is complex in nature and arises from diverse sources.<sup>17</sup> This means that a study registry modelled entirely on the structure of clinical trial registries is unlikely to be adequate, but this should not discourage efforts to develop and implement a prospective study registry suitable for public health. Retrospective assessment of outcome reporting bias as part of critical appraisal tools used in public health<sup>18 19</sup> is also possible but will require development of a tool such as ORBIT to be more suitable for the study designs and reporting found in evaluations of public health interventions.

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**REFERENCES**

- Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions 5.0.2* (Updated September 2009). The Cochrane Collaboration, 2009. [www.cochrane-handbook.org](http://www.cochrane-handbook.org) (accessed 13 May 2011).
- Hutton JL, Williamson PR. Bias in meta-analysis due to outcome variable selection within studies. *Appl Stat* 2000;**49**:359–70.
- Williamson PR, Gamble C. Identification and impact of outcome selection bias in meta-analysis. *Stat Med* 2005;**24**:1547–61.
- Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health* 2004;**94**:400–8.
- Smith RD, Petticrew M. Public health evaluation in the twenty-first century: time to see the wood as well as the trees. *J Public Health (Oxf)* 2010;**32**:2–7.
- Law C. Policy and evidence-based public health. In: Killoran A, Kelly MP, eds. *Evidence-Based Public Health: Effectiveness and Efficiency*. Oxford: Oxford University Press, 2010:16–26.
- Armstrong R, Waters E, Moore L, et al. Improving the reporting of public health intervention research: Advancing TREND and CONSORT. *J Public Health (Oxf)* 2008;**30**:103–9.
- Chan AW, Hrobjartsson A, Haahr MT, et al. Empirical evidence for selective reporting of outcomes in randomized trials—comparison of protocols to published articles. *JAMA* 2004;**291**:2457–65.

**What is already known on this subject**

Outcome reporting bias is an important issue in RCT literature and can lead to conclusions of overestimated effects.

**What this study adds**

- ▶ Outcome reporting bias exists in the public health literature.
- ▶ The exclusion of missing or incompletely reported outcomes can lead to a high risk of bias in a review and therefore the possible overestimation of intervention effectiveness.
- ▶ A prospective public health interventions study registry would aid the identification of unreported or incompletely reported outcomes.

9. **Hahn S**, Williamson PR, Hutton JL. Investigation of within-study selective reporting in clinical research: follow-up of applications submitted to a local research ethics committee. *J Eval Clin Prac* 2002;**8**:353–9.
10. **Kirkham JJ**, Dwan KM, Altman DG, *et al*. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;**340**:c365.
11. **Waters E**, Priest N, Armstrong R, *et al*. The role of a prospective public health intervention study register in building public health evidence: proposal for content and use. *J Public Health (Oxf)* 2007;**29**:322–7.
12. **Pearson M**, Garside R, Moxham T, *et al*. *Preventing Unintentional Injuries Among Under-15s in the Home. Report 1: Systematic Reviews of Effectiveness and Cost-Effectiveness of Home Safety Equipment and Risk Assessment Schemes*. Exeter: Peninsula Technology Assessment Group, 2009.
13. **Posner JC**, Hawkins LA, Garcia-Espana F, *et al*. A randomized, clinical trial of a home safety intervention based in a emergency department setting. *Pediatrics* 2004;**113**:1603–7.
14. **Paul CL**, Sanson-Fisher RW, Redman S, *et al*. Preventing accidental injury to young children in the home using volunteers. *Health Promot Int* 1994;**9**:241–9.
15. **Mallonee S**, Istre GR, Rosenberg M, *et al*. Surveillance and prevention of residential-fire injuries. *N Engl J Med* 1996;**335**:27–31.
16. **Mathieu S**, Boutron I, Moher D, *et al*. Comparison of registered and published primary outcomes in randomised controlled trials. *JAMA* 2009;**302**:977–84.
17. **Anderson LM**, McQueen DV. Informing public health policy with the best available evidence. In: Killoran A, Kelly MP, eds. *Evidence-Based Public Health: Effectiveness and Efficiency*. Oxford: Oxford University Press, 2010:437–47.
18. **Effective Public Health Practice Project**. *Quality Assessment Tool for Quantitative Studies*, 2009. <http://www.ehpp.ca/Tools.html> (accessed 13 May 2011).
19. **National Institute for Health & Clinical Excellence**. *Methods for the Development of NICE Public Health Guidance*. London: National Institute for Health & Clinical Excellence, 2009.

# Advancing Postgraduates. Enhancing Healthcare.

The *Postgraduate Medical Journal* is dedicated to advancing the understanding of postgraduate medical education and training.

- Acquire the necessary skills to deliver the highest possible standards of patient care
- Develop suitable training programmes for your trainees
- Maintain high standards after training ends

*Published on behalf of the fellowship for Postgraduate Medicine*

FOR MORE DETAILS OR TO SUBSCRIBE,  
VISIT THE WEBSITE TODAY

**postgradmedj.com**

ESSENTIAL  
READING FOR  
PLAB  
EXAMINEES



**BMJ | Journals**