



ORIGINAL ARTICLE

Survey of new 2007 and 2011 Cochrane reviews found 37% of prespecified outcomes not reported

Valerie Smith^{a,*}, Mike Clarke^b, Paula Williamson^c, Elizabeth Gargon^d

^aSchool of Nursing & Midwifery, Trinity College Dublin, 24 D'Olier Street, Dublin 2, Ireland

^bAll-Ireland Hub for Trials Methodology Research, Queens University Belfast, Grosvenor Road, Belfast BT12 6BA, Northern Ireland

^cMRC North West Hub for Trials Methodology Research, University of Liverpool, Liverpool, L69 3GS, UK

^dDepartment of Biostatistics, University of Liverpool, Liverpool, L69 3GA, UK

Accepted 24 September 2014; Published online xxxx

Abstract

Objectives: To survey the outcomes used in Cochrane Reviews, as part of our work within the Core Outcome Measures in Effectiveness Trials Initiative.

Study Design and Setting: A descriptive survey of Cochrane Reviews, divided by Cochrane Review Group (CRG), published in full for the first time in 2007 and 2011. Outcomes specified in the methods section of each review and outcomes reported in the results section of each review were of interest, in this exploration of the common use of outcomes and core outcome sets (COS).

Results: Seven hundred eighty-eight reviews, specifying 6,127 outcomes, were included. When we excluded specified outcomes from the 86 reviews that did not include any studies, we found that 1,996 (37%) specified outcomes were not reported. Of the 361 new reviews with studies from 2011, 113 (31%) had a “summary of findings” table (SoF). Fifteen broad outcome categories were identified and used to manage the outcome data. We found consistency in the use of these categories across CRGs but inconsistency in outcomes within these categories.

Conclusion: COS have been used rarely in Cochrane Reviews, but the introduction of SoF makes the development and application of COS timelier than ever. © 2014 Elsevier Inc. All rights reserved.

Keywords: Core outcome set; COS; COMET; Outcome measurement; Systematic reviews; Cochrane collaboration

1. Introduction

Cochrane Reviews, with over 5,500 published in full online by 2014, have been described as “unique because they are both produced by, and are relevant to, everyone interested in the effects of human health care” [1]. The Cochrane Collaboration, celebrating its 21st “birthday” this year, prides itself on preparing, maintaining, and promoting access to high-quality, timely research evidence

for health care decision making. It does this by supporting the rigorous conduct and reporting of systematic reviews. However, one of the difficulties often faced by systematic reviewers, when synthesizing the evidence from individual studies, is heterogeneity in the outcomes measured in these studies [2]. This difficulty not only presents when clinical trials on a similar topic or condition use a variety of outcome measures, but also where researchers measure the same outcome in a variety of ways [3]. Adding to this is the problem of selective reporting of outcomes in research reports, which can lead to outcome reporting bias [4–6]. This bias arises when a selection of the originally recorded outcomes is chosen for reporting in study publications, on the basis of their results [7]. The extent of this was highlighted in a review of 2,562 trials included in 283 Cochrane Reviews [8]. Outcome reporting bias was suspected in at least one trial in 35% of the examined reviews. In a sensitivity analysis of 81 of the included reviews that had a single meta-analysis of the review primary outcome, an assessment of the impact of outcome reporting bias was

Conflict of interest: One of the authors (V.S.) was a research fellow (part-time, 0.5 full-time equivalent), under the FP7 award, to conduct work for the COMET Initiative. The survey reported here was performed within this employment contract. Although authors V.S., M.C., and P.W. are members of the Cochrane Collaboration, the views expressed in this article are not necessarily those of the Collaboration or the Cochrane groups with which they are associated.

Funding: The COMET Initiative is pleased to acknowledge funding from the European Union Seventh Framework Programme ([FP7/2007–2013] [FP7/2007–2011]) under grant agreement n°305081.

* Corresponding author. Tel.: 00353-1-8964031; fax: 00353-1-8963001.

E-mail address: smithv1@tcd.ie (V. Smith).

What is new?**Key findings**

- Nonreport of specified outcomes in Cochrane Reviews is considerable (37%).
- The median figure for specified outcomes in Cochrane Reviews is 7, but many reviews (>20%) specify 10 outcomes or more.

What this adds to what was known?

- There is wide variation in outcomes specified in Cochrane Reviews.
- Bias in reporting of specified outcomes might be a problem in some systematic reviews.

What is the implication and what should change now?

- Outcome variation and outcome reporting bias may be addressed by developing and applying agreed standardized sets of outcomes, known as “core outcome sets.”

performed. The analysis demonstrated that 52 of the 81 reviews included at least one trial that had a high suspicion for outcome reporting bias and the treatment effect estimate was reduced by 20% or more in 23% ($n = 19$) of the reviews [8]. This indicates that outcome reporting bias can present a substantial problem for those who wish to use the findings of Cochrane Reviews, other systematic reviews, and individual trials themselves when making health care decisions.

One way to address the difficulty of outcome variation and outcome reporting bias in systematic reviews is to develop and apply agreed standardized sets of outcomes, known as “core outcome sets” (COS) [2–4,9]. For example, published reports on COS are available for asthma in children [10], ulcerative colitis [11], models of maternity care [12], and, most noticeably, rheumatology [13,14]. The Core Outcome Measures in Effectiveness Trials (COMET) Initiative, launched in 2010, is further advancing efforts for COS development [2,4,9]. This Initiative brings researchers interested in COS together to consider methods for COS development and to highlight health care areas in need of COS. The Initiative supports the idea that a COS should represent the minimum to be measured and reported in all clinical trials on a specific condition, while recognizing that outcomes outside the COS might also be important in the context of each study. This use of the COS as a minimum across an entire research area would allow for the results of trials and other studies to be effectively compared, contrasted, and combined, as appropriate [2,9].

One of the objectives of the COMET Initiative is to develop a strategy to link the development of COS for trials with the specification of outcomes for Cochrane Reviews, including the outcomes in their summary of findings (SoF) table. Developed with the GRADE (Grading of Recommendations Assessment, Development and Evaluation) group, SoF tables have been possible in Cochrane Reviews since 2008. A SoF table presents the main findings of a Cochrane Review in a simple tabular format. These tables provide information on the quality of evidence and on the magnitude of effect of the interventions examined in a review. They allow for the inclusion of up to seven important reported outcomes, providing a way to present the main findings of Cochrane Reviews in a simple and transparent format [15]. They have been shown to help readers understand the results of Cochrane Reviews more correctly and faster and are considered to facilitate a more effective and efficient uptake of key information [16]. However, the SoF table will only be effective for evidence transfer if the outcomes selected for inclusion in the table are appropriate to the review question. To explore these issues as part of our work within COMET, we have performed a survey of Cochrane Reviews to identify the variety of outcome measures used in them.

2. Aim and objectives

The aim of the study was to survey the outcomes used in Cochrane Reviews. The objectives of the survey were as follows:

1. To identify the variety of outcome measures used in Cochrane Reviews.
2. To identify and highlight the use of COS in reviews from Cochrane Review Groups (CRG).
3. To identify health care areas that might benefit from COS development.

3. Methods*3.1. Design*

A descriptive survey of Cochrane Reviews, divided by CRG, published in full for the first time in 2007 and 2011, was performed. Newly published 2007 and 2011 reviews were purposively chosen so as to explore any potential change in outcome specifying and reporting over time. Outcomes specified in the methods section of each review and outcomes reported (defined as a reported result on an outcome in the text of the review) in the results section of each review were of interest. We also evaluated the 2011 reviews for their use of “summary of findings” tables. The survey was conducted between November 2012 and April 2013.

3.2. Data extraction and management

Cochrane Reviews that were published in full for the first time in 2007 or 2011 were identified from the Cochrane

Collaboration's internal database, Archie. Full texts of these reviews were then obtained from the Cochrane Database of Systematic Reviews in The Cochrane Library, accessing the appropriate "earlier version" if the review had been updated since its original publication in 2007 or 2011. Data extraction tables were predesigned and used to extract, store, and manage the relevant data. Separate data extraction tables were built for each CRG. The extracted data allowed us to determine the number of newly published reviews in 2007 and 2011 for each CRG, number of trials included in each review, number of participants included in each review, outcomes specified in the methods section of each review, and outcomes reported in the results section of each review.

3.3. Data analysis

Data analysis involved an in-depth exploration and description of specified and reported outcomes in the new reviews from each CRG.

4. Findings

Three hundred eighty-seven newly published Cochrane Reviews in 2007 from 47 CRGs and 401 newly published Cochrane Reviews in 2011 from 50 CRGs were identified and included. The total number of outcomes specified in the methods sections of all these new reviews was 6,127. When specified outcomes from reviews that had not included any studies were excluded ($n = 86$ reviews), the proportion of specified outcomes that was reported was 63% (3,367 of 5,363). Of the nonreported specified outcomes, 23% (1,264) were not reported because of nonmeasurement of the outcome in the included studies or insufficient data to report. However, for 14% (732) of outcomes, we could not find a reason in the text of the review for why the outcome was not reported.

The number of specified outcomes in each review was assessed. These ranged between 1 (10 reviews) and 26 (1

review) outcomes (Fig. 1). The median figure was seven outcomes.

A scoping overview involving a careful check of the extracted data to determine the extent, range, and nature of the outcomes identified fifteen broad categories of outcomes that emerged prominently across the CRGs. Rather than attempting to define outcome domain systems, as others have done [10,14,17], we used these categories to extract and manage the data for analyses purposes in this survey. "Mapping" these category outcomes to other systems, such as that recently described by Boers et al. [14], is likely to be possible, and this is addressed in the discussion section of this article. The broad categories of outcomes used in our survey encompass person-level outcomes, resource-based outcomes, and research/study-related outcomes. They are adverse events or effects (AE), mortality/survival, infection, pain, other physiological or clinical, psychosocial, quality of life, activities of daily living (ADL), medication, economic, hospital, operative, compliance (with treatment), withdrawal (from treatment or study), and satisfaction (patient, clinician, or other health care provider).

Table 1 illustrates this data management process including outcome examples within each category, for one CRG.

Table 2 provides examples of specified outcome variation/consistency in nine selected outcome categories as follows: "adverse events/effects," "infection," "medication," "mortality/survival," "pain," "quality of life," "economic," "hospital," and "compliance," in 2007 and 2011. Although the survey found greater consistency in the outcome categories across the CRGs, the individual outcome measures were much less consistent within some categories. For example, in 2007, outcomes in the category "hospital" were specified in reviews from 32 of the 47 CRGs (68%). A total of 84 "hospital" outcomes were specified in these 32 reviews and included 23 individual outcome types. (Table 2 provides an example of five of these individual "hospital" outcome types.) In addition,

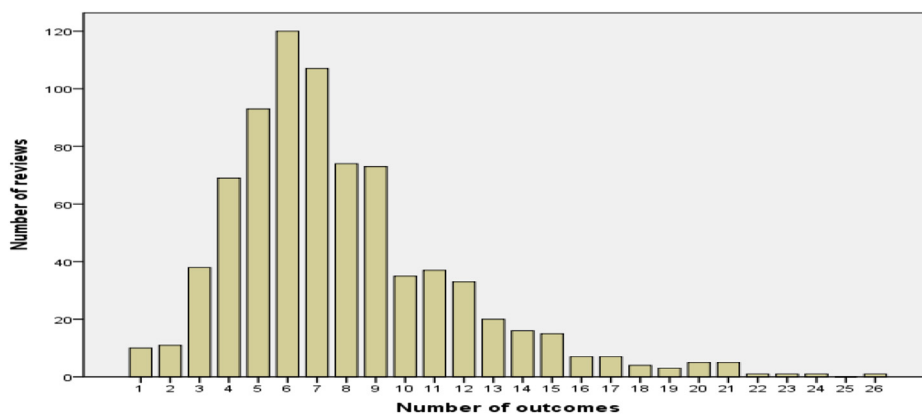


Fig. 1. Number of outcomes in number of reviews.

Table 1. Example of CRG outcome data management table (shortened for illustrative purposes)

Pregnancy and childbirth group 2011 (32 reviews)			
Category	Outcomes	Number of outcomes specified	Number of outcomes reported
Adverse events/effects	AE: type specified	7	2 ¹
	AE: any	10	4 ³
Mortality/survival	Mortality: all cause (maternal)	8	2 ³
	Stillbirth	7	1 ²
Infection	Mortality: all cause (neonatal)	10	2 ⁵
	Infection: any/type not specified	15	3 ⁴
	Pyrexia	1	0
Pain	Neonatal sepsis	4	1
	Pain: any/type unspecified	10	9
Physiological or clinical	Pain: relief/control of	4	4
	Blood loss	5	2 ¹
	Breast feeding	11	3 ¹
	Instrumental birth	11	7 ²
	Apgar <7 at 5 minutes	13	10 ¹
Psychosocial	Gestation at birth	7	4 ³
	Mother–baby interaction	5	0
	Perceived control in labor	5	1
	Anxiety	4	4
Quality of life	Depression	2	1
	Quality of life generic	2	0
ADL	Sleep	1	0
	Interference with ADL	1	0
Medication	Hypertensive drugs	1	0
	Use of magnesium sulfate	1	0
Economic	Cost	8	1 ²
	Cost-effectiveness	1	0
	Health care resource use	2	0 ¹
Hospital	Length of hospital stay	6	1 ²
	Readmission to hospital	1	0
	Admission to NICU	15	9 ³
Operative	Surgical intervention	1	0
	Duration of operation	1	1
Compliance	Compliance with treatment	1	0
Withdrawal	Withdrawal from treatment	1	0
Satisfaction	Maternal satisfaction	13	4 ²
	Provider satisfaction	1	0

Abbreviations: CRG, Cochrane Review Group; AE, adverse events or effects; ADL, activities of daily living; NICU, Neonatal Intensive Care Unit. ⁿ indicates number of specified outcomes not reported due to 0 studies in reviews.

of these 23 individual outcome types, some appeared more often in the reviews of multiple CRGs than others. For example, “length of hospital stay” was used in at least one review by 81% of the CRGs. This can be contrasted to the other “hospital” outcomes, such as “admission to ICU” or “emergency department visit,” which were used by 19% and 17% of CRGs, respectively. Outcome variation for 2011 is illustrated by the category “mortality/survival.” Outcomes in this category were present in reviews from 35 of the 50 CRGs (70%). A total of 103 “mortality/survival” outcomes were specified in these 35 reviews and included 27 individual outcome types (Table 2 provides an example of five of these individual “mortality/survival” outcome types). “All-cause mortality” was the outcome type most frequently used in this category (appearing in at least one review in 54% of the CRGs). This can be contrasted to other “mortality/survival” outcomes such as “overall survival” or “progression-free survival,” which were used in 20% and 9% of CRGs, respectively.

The outcome category “physiological or clinical” was consistent across all the CRGs in 2007 and 2011; that is, at least one review in all CRGs specified at least one outcome in this category, which is not surprising given the CRGs’ focus on health care. This category contained the largest number of outcomes in this survey and the greatest outcome variation. This was because most of the outcomes in this category were related to diverse physical bodily functioning and/or multiple body systems (eg, cardiovascular respiratory, neurological, reproductive, renal, hepatic and so forth). For this reason, “physiological or clinical” outcomes tended, in the main, to be review and/or CRG specific. To illustrate outcome variation within this category and across the 2007 and 2011 reviews within a CRG, Table 3 provides examples of “physiological or clinical” outcomes for four selected CRGs.

The proportion of specified outcomes reported after removing reviews that did not include any studies was calculated for each CRG. Fig. 2 and Fig. 3 present the

Table 2. Category outcome variation/consistency (specified outcomes only; shortened for illustrative purposes)

Outcome category	Individual outcome	No of CRGs with outcome in ≥ 1 review/no CRGs using outcome category (%) 2007	No of CRGs with outcome in ≥ 1 review/no CRGs using outcome category (%) 2011
Adverse events/effects	AE (any)	43/46 (93)	46/48 (96)
	AE (type specific)	20/46 (43)	22/48 (46)
	Complications (any)	14/46 (30)	10/48 (21)
	Complications (type specific)	7/46 (15)	5/48 (10)
Infection	Infection (type specific)	12/23 (52)	7/20 (35)
	Infection (any)	10/23 (43)	9/20 (45)
	Duration of infection	3/23 (13)	—
	Fever	3/23 (13)	1/20 (5)
	Fever/parasitological clearance	2/23 (1)	1/20 (5)
	Medication	Consumption (type specific)	13/24 (54)
Medication	Need for rescue medication	6/24 (25)	8/18 (44)
	Change to medication	3/24 (13)	—
	Prescription rates	—	2/18 (11)
	Medication-free days	—	1/18 (6)
Mortality/survival	Mortality (all cause)	13/33 (39)	19/35 (54)
	Mortality (cause specific)	17/33 (52)	17/35 (49)
	Survival (overall)	10/33 (30)	7/35 (20)
	Survival (disease free)	5/33 (15)	4/35 (11)
	Survival (progression free)	3/33 (9)	3/35 (9)
Pain	Pain (any)	16/27 (59)	12/21 (57)
	Pain (type specific)	9/27 (33)	8/21 (38)
	Pain relief/control	7/27 (26)	12/21 (57)
	Pain intensity	7/27 (26)	7/21 (33)
	Pain score	6/27 (22)	4/21 (19)
	Duration of pain	3/27 (11)	—
	Quality of life	Quality of life (generic)	38/38 (100)
Economic	Quality of life (disease specific)	7/38 (18)	6/38 (16)
	Costs	19/31 (61)	22/30 (73)
	Health care resource use	13/31 (42)	11/30 (37)
	Cost-effectiveness	5/31 (16)	10/30 (33)
Hospital	Economic data	2/31 (1)	2/30 (1)
	Length of hospital stay	26/32 (81)	23/28 (82)
	Admission to hospital	15/32 (47)	14/28 (50)
	Readmission to hospital	8/32 (25)	5/28 (18)
	Admission to ICU/NICU	6/32 (19)	2/28 (7)
	Emergency department visit	5/32 (17)	2/28 (7)
Compliance	Treatment compliance	12/18 (67)	6/14 (43)
	Adherence to treatment	4/18 (22)	5/14 (36)
	Attendance at course	2/18 (11)	—
	Compliance measures	1/18 (1)	—

Abbreviations: CRG, Cochrane Review Group; AE, adverse events or effects.

results, combined for 2007 and 2011, respectively, for all outcomes and primary outcomes only.

Turning to the SoFs in the 2011 reviews, we found that 31% (113 of 361) of reviews that had studies included in them, presented a SoF. Of the 113 reviews, 27% ($n = 31$) included the maximum of seven outcomes in their SoF, 16% ($n = 18$) included six outcomes, 11% ($n = 12$) included five outcomes, 13% ($n = 15$) included four outcomes, 11% ($n = 12$) included three outcomes, 7% ($n = 8$) included two outcomes, and 15% ($n = 17$) of reviews included one outcome.

A final exercise in this survey was to identify individual outcomes, which were used in more than half of the included reviews for each CRG. This identified commonly listed outcomes within each CRG and might help identify the current use of COS by CRGs or areas where COS development might be targeted. [Additional File 1 \(Appendix at](#)

www.jclinepi.com) illustrates this for the 2007 and 2011 newly published reviews.

5. Discussion

In this survey, outcomes used in the 788 Cochrane Reviews published in full for the first time in 2007 and 2011 were explored. A total of 6,127 specified outcomes were listed in the methods sections of the included reviews. After excluding reviews that did not include any studies, the results demonstrated that 37% of specified outcomes in the newly published Cochrane Reviews were not reported. Legitimate reasons for nonreport in the review (ie, nonmeasurement of the outcome in the included studies or insufficient data to report) were provided in the text of the review for 23% of cases, but we found no explanations,

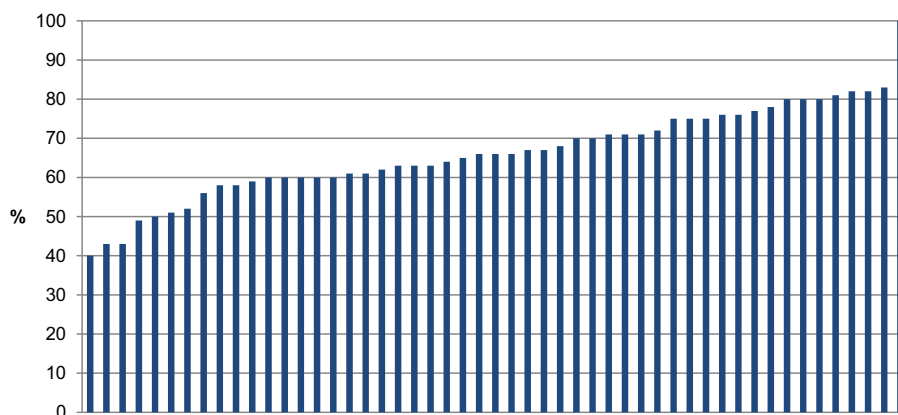
Table 3. Examples of “physiological or clinical” outcomes in four selected CRGs (shortened for illustrative purposes)

CRG	2007	2011
Acute respiratory infections	Exacerbations Lung function FEV1 Biomarkers/airway markers Exercise capacity Severity of symptoms PEF rate	Clinical improvement Relapse/recurrence Requiring ventilator support Change in symptoms Oxygenation level Intubation rate Evidence of myocardial injury
Heart	Thromboembolism Aneurysm Ventricular fibrillation Heart failure Hypotension Cardiovascular events: composite MI Stroke	MI Heart failure Stroke Thrombocytopenia Return of circulation Cardiovascular events: composite Blood transfusion Unstable angina
Oral health	Degree of function Aesthetics Inflammation Nerve injury No. fillings/crowns retained Mandibular length Open bite correction	Lateral movement and protrusion Damage to teeth Decay Gingival health Filled teeth Xerostomia Salivary flow
Wounds	Time to healing No. of wounds healed in specific time Wound hematoma Wound dehiscence Fungating tumor containment Exudate Malodor Hemorrhage	Scarring Time to healing No. of wounds healed in specific time Amputations Recurrence Change in wound area Cosmetic appearance Ease of dressing removal

Abbreviation: CRG, Cochrane Review Group; FEV1, forced expiratory volume exhaled at the end of the first second of forced expiration; PEF, peak expiratory flow; MI, myocardial infarction.

in the text of the review, for the other 14%. Furthermore, when the proportions of specified and reported outcomes, from reviews within each CRG, were explored (Fig. 2), the results revealed that the proportion of specified outcomes that were reported ranged from 40% to 100% and was between 50% and 75% for 30 CRGs (60%) and 75% or higher for 14 CRGs (28%). A median of seven outcomes across all included reviews was identified, but many

included reviews (> 150) specified 10 outcomes or more (Fig. 1). Acknowledging that a definitive answer as to “how many” outcomes should be specified in a review is unlikely to be achievable, specifying too many outcomes may still present difficulties for reviewers in identifying all outcome data for reporting in their reviews. This suggests, perhaps, that greater consideration of outcome selection may be required during review development.

**Fig. 2.** Proportion of specified outcomes reported by Cochrane Review Group: all outcomes.

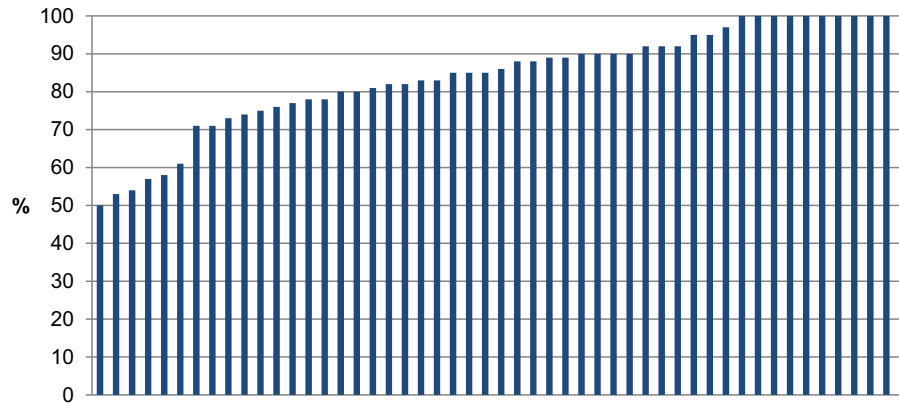


Fig. 3. Proportion of specified outcomes reported by Cochrane Review Group: primary outcomes only.

There is sufficient evidence to suggest that the nonreport of specified outcomes in study publications and systematic reviews leads to bias [5,8] and even more so when the reasons for nonreport are not provided. Nonreporting of outcomes can result in a reduction in the quality of the presented evidence and can cause those accessing this evidence to interpret it inaccurately. If the findings of a systematic review are to influence clinical practice effectively, potential biases and the limitations of a review need to be made clear and need to be transparently stated. This allows the reader, provided with all the information, to consider the strength of the review's findings and how these may or may not be used in making a decision. Previous studies with trial researchers have found common reasons for nonreport of specified trial outcomes. These are a lack of statistical significance, a lack of understanding about the importance of reporting "negative" results, missing or delay in obtaining data, data perceived as being uninteresting, too few events to consider them worth reporting, and restrictions placed on reporting due to Journal space [6,18]. It is unclear from our survey, however, as to why explanations for the nonreport of many specified outcomes (14%) were not provided. By highlighting this finding, it is hoped that review authors will be alerted to the importance of explaining why they do not report specified outcomes in their review. Furthermore, peer reviewers of Cochrane Reviews and the editorial teams of CRGs might be encouraged to insist on this before recommending a review for publication.

The use of specific outcomes in at least half of the new reviews for a CRG was explored [Additional File 1 (Appendix at www.jclinepi.com)] to highlight the possible use of COS and to identify potential areas that might benefit from COS development. Using the findings from the 2011 reviews for discussion purposes, 41 of the 50 CRGs had at least one outcome that was specified by more than half its reviews. This finding might suggest that a COS that would include such outcomes might be appropriate for reviews relevant to that CRG. However, one needs to be cautious because the number of newly published reviews in some CRGs is low and some CRGs have a wide scope

making it unlikely that a single COS would be suitable across all their reviews. For example, in 2011, the anesthesia, back, depression, anxiety and neurosis, hypertension, schizophrenia, and skin CRGs each had four or fewer newly published reviews. Outcomes thus highlighted as being reported in at least half of these CRGs' reviews are outcomes reported in two, three, or four reviews only. It is plausible that these highlighted outcomes may not be used or necessarily applicable in a significantly high proportion of reviews, if all the reviews in these CRGs were considered. Of interest, our analysis demonstrated that 36 of the 51 CRGs (71%) in the 2011 survey used the outcome category "adverse events/effects" in more than half their reviews, indicating that not all reviews in CRGs are using this outcome. The Cochrane Handbook of Systematic Reviews of Healthcare Interventions [15] suggests that every health care intervention comes with some risk of harmful or AEs and recommends that all reviews should try to include some consideration of the adverse aspects of the interventions (section 14.1.1, www.handbook.cochrane.org).

The associated difficulties with outcome variation in synthesizing evidence from individual studies have been identified in previous research. For example, Green et al. [19] identified wide variation in clinical trials investigating the efficacy of interventions for shoulder pain. Blackwood et al. [20] highlight a lack of consistency in the measurement of ventilation outcomes and advocate the need for a minimum COS for trials involving mechanical ventilation. The COMET database [9] further provides examples of systematic reviews of outcomes that demonstrate variation. Our survey also revealed additional evidence of outcome variation. Although 15 broad categories of outcomes were identified as being used in reviews across CRGs, there was inconsistency in the description of the outcomes within these categories. For example, in 2007, outcome measurement related to clinical visits included, among others, appointment length, length of therapy, and time with therapist. Similarly, emergency department—related outcomes were measured as a dichotomous outcome in some cases (eg, visit to the department) but as a continuous outcome

in others (eg, length of visit and mean reduction in visit). Variation in both the description and the way an outcome is measured can present difficulties for health care providers who wish to use this evidence in the context of their health care provision or service. This is made even more problematic because the outcome might be measured by different instruments that are not comparable, the reliability and validity of different instruments might vary, and the measurements may be completed by study populations at varying times [3]. Standardizing how an outcome should be measured is essential if the development and implementation of COS is to be successful [2]. Although work in this area is underway [see www.cosmin.nl, the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) Initiative], further work is required.

The 15 broad categories of outcomes used in our survey emerged from an initial scoping exercise of all outcomes in all included reviews. In using these categories, we did not attempt to define outcome domains, but, rather, we sought to develop a method to manage and organize the outcome data. Boers et al., [14] as part of their work in Outcome Measures in Rheumatology (OMERACT), recently published an upgrade of the OMERACT framework and consider this as a potential template for the development of COS in other subspecialties. The framework describes four “core areas” within which there are specific domains. These areas are death, life impact, resource use/economic impact, and pathophysiological manifestations. AEs, although labeled separately, are measured within each of the core areas. The 15 broad outcome categories that emerged for use in our survey might, in retrospect, be collapsed further and “mapped” to these core areas. For example, our category of “mortality/survival” could be considered within the core area “death.” “Quality of life,” “activities of daily living,” “satisfaction,” and “psychosocial” could be aligned with OMERACTs “life impact.” “Infection,” “pain,” and “other physiological or clinical” could be aligned with “pathophysiological manifestations.” Finally, our categories of “medication,” “hospital,” “economic,” and “operative” could be aligned with “resource use/economic impact.”

The use of COS is advantageous for facilitating effective pooling of data from different studies on a specific condition, for meaningful comparisons of treatments and their effects, and for encouraging a more complete reporting of outcomes in studies [2,21,22]. The Coordinating Editors of CRGs also appear to support the concept of using COS for outcome selection in systematic reviews. In a survey of these Editors [23], 87% ($n = 39$) responded that standardizing all outcomes across all reviews for a specific condition would be advantageous for systematic reviews and meta-analyses. Thirty-three (73%) also said that a COS for effectiveness trials should be used in a SoF table. The reasons given for this varied and included the inclusion of relevant outcomes (30%), reduction in bias (3%), comparability of outcomes (18%), and improved interpretability of outcomes (12%). This support for COS use in SoFs by Coordinating Editors

contrasts with data in this survey, where a 31% SOF use was identified in the 2011 reviews.

In developing and applying a COS for specific topics and interventions, it is important also to consider COS use in context. For example, a developed COS for studies on human immunodeficiency virus may include death as an important core outcome. However, with advances in health care, biomedicine, and pharmacology, death as an outcome might become less important to this topic over time. Alternative outcomes, such as side effects and adherence to treatment, for example, may emerge as more relevant current outcomes. For this reason, to avoid stasis in developed COS, COS should be viewed as potentially dynamic and evolutionary, requiring revisiting and updating every few years. Furthermore, it is important to remember that the developed COS should be considered as the minimum to be measured and reported for trials in a specific condition; if outcomes outside this COS are relevant and important in the context of an individual study, they should also be measured and reported. This may emerge even more so in conditions where composite outcome measures might be used. Consideration also needs to be given to how outcomes are measured, but this is beyond the scope of our current work, where the focus is on what to measure.

The survey reported here explored only one aspect of outcome measurement in Cochrane Reviews, that is the type and variety of outcomes measured. Further research on the topic of outcome measurement in Cochrane Reviews is required. Current plans include a repeat survey for Cochrane Reviews published in full for the first time in 2013, to extend our assessment of whether there have been changes in the specification and reporting of outcomes over time.

6. Conclusion

This survey suggests that there has been minimal use and implementation of COS in Cochrane Reviews. It also reveals, for a large proportion of outcomes, that a reason for not reporting a specified outcome is not provided in the text of a review. This indicates that outcome reporting bias might be a problem for some systematic reviews. With the introduction of SoF tables in Cochrane Reviews, work to develop and apply COS is timelier than ever. Developing and implementing COS is not without its challenges, however. The COMET Initiative, through collaboration, stakeholder involvement, consensus, review, and feedback, is committed to advancing methods for rigorous COS development and implementation. However, agreement, guidelines, and adherence will be required for the successful use of COS.

Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.jclinepi.2014.09.022>.

References

- [1] The Cochrane Collaboration Web site. Available at <http://www.cochrane.org/about-us/>. Accessed July 10, 2013.
- [2] Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13:132.
- [3] Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007;8:39.
- [4] Williamson PR, Clarke M. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative: its role in improving Cochrane reviews [editorial]. *Cochrane Database Syst Rev* 2012;5:ED000041.
- [5] Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A, Cronin E, et al. Systematic review of empirical evidence of study publication bias and outcome reporting bias. *PLoS One* 2008;3(8):e3081.
- [6] Smyth RMD, Kirkham JJ, Jacoby A, Altman DG, Gamble C, Williamson PR. Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. *Br Med J* 2010;341:C7153.
- [7] Hutton JL, Williamson PR. Bias in meta-analysis due to outcome variable selection within studies. *Appl Stat* 2000;49:359–70.
- [8] Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, Williamson PR. The impact of outcome reporting bias in randomized controlled trials on a cohort of systematic reviews. *Br Med J* 2010;340:C365.
- [9] The COMET Initiative website. Available at <http://www.comet-initiative.org/>. Accessed July 10, 2013.
- [10] Sinha IP, Gallagher R, Williamson PR, Smyth R. Development of a core outcome set for clinical trials in childhood asthma: a survey of clinicians, parents, and young people. *Trials* 2012;13:103.
- [11] Cooney RM, Warren BF, Altman DG, Abreu MT, Travis SP. Outcome measurement in clinical trials for ulcerative colitis: towards standardisation. *Trials* 2007;8:17.
- [12] Devane D, Begley CM, Clarke M, Horey D, O'Boyle C. Evaluating maternity care: a core set of outcome measures. *Birth* 2007;34(2):164–72.
- [13] Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 2007;8:38.
- [14] Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino M, et al. Developing core outcome measurement sets for clinical trials: OMERACT Filter 2.0. *J Clin Epidemiol* 2014;67:745–53.
- [15] Higgins JP, Green S: The Cochrane handbook for systematic reviews. 2011; Available at <http://handbook.cochrane.org/>. Accessed July 15, 2013.
- [16] Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. *J Clin Epidemiol* 2010;63:620–6.
- [17] World Health Organisation: International Classification of Functioning, Disability and Health (ICF). Available at <http://www.who.int/classifications/icf/en>. Accessed November 30, 2013.
- [18] Chan AW, Hrobjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols in published articles. *JAMA* 2004;291:2457–65.
- [19] Green S, Buchbinder R, Glazier R, Forbes A. Systematic review of randomized trials of interventions for painful shoulder: selection criteria, outcome assessment, and efficacy. *BMJ* 1998;316:354–60.
- [20] Blackwood B, Clarke M, McAuley DF, McGuigan PJ, Marshall JC, Rose L. How outcomes are defined in clinical trials of mechanically ventilated adults and children. *Am J Respir Crit Care Med* 2014;189:886–93.
- [21] Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003;106(3):337–45.
- [22] Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008;9:105–21.
- [23] Kirkham JJ, Gargon E, Clarke M, Williamson P. Can a core outcome set improve the quality of systematic reviews? A survey of the Coordinating Editors of Cochrane review groups. *Trials* 2013;14:21.