Journal of Clinical Epidemiology 81 (2017) 3−12

Journal of Clinical Epidemiology

REVIEWS

Authors seldom report the most patient-important outcomes and absolute effect measures in systematic review abstracts

Arnav Agarwal[a,b,1], Bradley C. Johnston[b,c,d,e,*,1], Robin W.M. Vernooij[f], Alonso Carrasco-Labra[b,g], Romina Brignardello-Petersen[b,g], Ignacio Neumann[b,h], Elie A. Akl[b,i], Xin Sun[b,j], Matthias Briel[b,k], Jason W. Busse[b,l,m], Shanil Ebrahim[b,m], Carlos E. Granados[b,n], Alfonso Iorio[b], Affan Irfan[o], Laura Martínez García[f], Reem A. Mustafa[b,p], Anggie Ramirez-Morera[q], Anna Selva[f], Ivan Solà[f], Andrea J. Sanabrai[f], Kari A.O. Tikkinen[r], Per O. Vandvik[s], Yuqing Zhang[b], Oscar E. Zazueta[f], Qi Zhou[b], Holger J. Schunemann[b], Gordon H. Guyatt[b], Pablo Alonso-Coello[b,f,t]

[a]*School of Medicine, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada*
[b]*Department of Clinical Epidemiology & Biostatistics, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada*
[c]*Institute of Health Policy, Management & Evaluation, University of Toronto, 4th Floor, 155 College St, Toronto, Ontario M5T 3M6, Canada*
[d]*Systematic Overviews through advancing Research Technology (SORT), Child Health Evaluative Sciences, The Hospital for Sick Children Research Institute, University of Toronto, 686 Bay Street, Toronto, Ontario M5G 0A4, Canada*
[e]*Department of Anesthesia & Pain Medicine, The Hospital for Sick Children, 555 University Ave., Toronto, Ontario M5G 1X8, Canada*
[f]*Clinical Epidemiology and Public Health Department, Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), C/ Sant Antoni Maria Claret, 167, Pavelló 18, planta 0, 08025 Barcelona, Spain*
[g]*Evidence-Based Dentistry Unit, Faculty of Dentistry, Universidad de Chile, Sergio Livingstone 943, Santiago, Chile*
[h]*Department of Medicine, Pontificia Universidad Católica de Chile, Avenida Libertador Bernardo O Higgins 340, Santiago, Región Metropolitana, Chile*
[i]*Department of Internal Medicine, American University of Beirut Medical Center, P.O. Box: 11-0236, Riad-El-Solh, Beirut 1107 2020, Lebanon*
[j]*Chinese Evidence-based Medicine Center, West China Hospital, Sichuan University, 37 Guo Xue Xiang, Chengdu 610041, China*
[k]*Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital Basel, Spitalstrasse 12, Basel 4031, Switzerland*
[l]*Michael G. DeGroote Institute for Pain Research and Care, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada*
[m]*Department of Anesthesia, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada*
[n]*Área de investigaciones, Facultad de Medicina, Universidad de La Sabana, Campus del Puente del Común Km, 7 Autopista Norte, Chía, Colombia*
[o]*Internal Medicine Residency Program, University of Illinois, Urbana-Champaign, 506 S. Mathews Ave., Urbana, IL 61801, USA*
[p]*Departments of Medicine and Biomedical & Health Informatics, University of Missouri-Kansas City, School of Medicine, M4-303, 2411 Holmes St, Kansas City, MO, USA*
[q]*CCSS Permanent Medical Advisor, Health Care Development Division, IHCAI Foundation & Central America Cochrane, 1st Ave., 35th and 37th St, Number 3530, Barrio Escalante, San José, Costa Rica*
[r]*Departments of Urology and Public Health, Helsinki University Hospital and University of Helsinki, Haartmaninkatu 4, Helsinki 00029, Finland*
[s]*Norwegian Knowledge Centre for the Health Services, P.O. Box 4404 Nydalen, Oslo 0403 Norway*
[t]*CIBER of Epidemiology and Public Health (CIBERESP-IIB Sant Pau), Av. Monforte de Lemos, 3-5, Pabellón 11, Planta 0, Madrid 28029, Spain*

Accepted 9 August 2016; Published online 21 August 2016

Competing interests: All authors have completed the International Committee of Medical Journal Editors (ICMJE) uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare no support from any organization for the submitted; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Contributions of all authors: *Conception and design*: A.A., B.C.J., R.B.-P., and P.A.-C. *Design of search strategy*: A.C.-L. and P.A.-C. *Paper selection*: B.C.J., R.W.M.V., P.A.-C., A.C.-L., R.B.-P., I.N., E.A.A., X.S., M.B., J.W.B., S.E., C.E.G., A.I., Af.I., L.M.G., R.A.M., A.R.-M., A.S., I.S., A.J.S., K.A.O.T., P.O.V., Y.Z., O.E.Z., and P.A.-C. *Data abstraction*: B.C.J., R.W.M.V., P.A.-C., A.C.-L., R.B.-P., I.N., E.A.A., X.S., M.B., J.W.B., S.E., C.E.G., A.I., Af.I., L.M.G., R.A.M., A.R.-M., A.S., I.S., A.J.S., K.A.O.T., P.O.V., Y.Z., O.E.Z., and P.A.-C. *Data analysis*: Q.Z., A.A., and R.B.-P. *Interpretation of results*: A.A., B.C.J., R.W.M.V., P.A.-C., A.C.-L., R.B.-P., I.N., E.A.A., X.S., M.B., J.W.B., S.E., C.E.G., A.I., Af.I., L.M.G., R.A.M., A.R.-M., A.S., I.S., A.J.S., K.A.O.T., P.O.V., Y.Z., O.E.Z., H.J.S., G.H.G., and P.A.-C. *Manuscript drafting*: A.A., B.C.J., and G.H.G. *Manuscript review and approval*: A.A., B.C.J., R.W.M.V., P.A.-C., A.C.-L., R.B.-P., I.N., E.A.A., X.S., M.B., J.W.B., S.E., C.E.G., A.I., Af.I., L.M.G., R.A.M., A.R.-M., A.S., I.S., A.J.S., K.A.O.T., P.O.V., Y.Z., O.E.Z., H.J.S., G.H.G., and P.A.-C.

[1] The first two authors contributed equally to this work and are both considered as first authors.

* Corresponding author. Tel.: +1-416-813-7654x328524; fax: 416-813-7543.

E-mail address: bradley.johnston@sickkids.ca (B.C. Johnston).

http://dx.doi.org/10.1016/j.jclinepi.2016.08.004
0895-4356/© 2016 Elsevier Inc. All rights reserved.

Abstract

Objectives: Explicit reporting of absolute measures is important to ensure treatment effects are correctly interpreted. We examined the extent to which authors report absolute effects for patient-important outcomes in abstracts of systematic review (SR).

Study Design and Setting: We searched OVID MEDLINE and Cochrane Database of Systematic Reviews to identify eligible SRs published in the year 2010. Citations were stratified into Cochrane and non-Cochrane reviews, with repeated random sampling in a 1:1 ratio. Paired reviewers screened articles and recorded abstract characteristics, including reporting of effect measures for the most patient-important outcomes of benefit and harm.

Results: We included 96 Cochrane and 94 non-Cochrane reviews. About 117 (77.5%) relative measures were reported in abstracts for outcomes of benefit, whereas only 34 (22.5%) absolute measures were reported. Similarly, for outcomes of harm, 41 (87.2%) relative measures were provided in abstracts, compared with only 6 (12.8%) absolute measures. Eighteen (9.5%) abstracts reported both absolute and relative measures for outcomes of benefit, whereas only two (1.1%) abstracts reported both measures for outcomes of harm. Results were similar between Cochrane and non-Cochrane reviews.

Conclusion: SR abstracts seldom report measures of absolute effect. Journal editors should insist that authors report both relative and absolute effects for patient-important outcomes. © 2016 Elsevier Inc. All rights reserved.

Keywords: Absolute measures; Patient-important outcomes; Explicit reporting; Abstract reporting; Cochrane reviews; Non-Cochrane reviews

1. Introduction

Readers of the medical literature often first refer the abstract of a systematic review (SR) to judge the magnitude of results, including the point estimates for benefit and harm and their associated confidence intervals [1], and may, because of limited access to full-text journals and time constraints, rely solely on abstracts. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines for abstracts recommend that results for the main outcomes of both benefits and harms should be reported as part of the synthesis of the results, preferably providing the number of studies and participants for each included outcome, summary measures (estimates), and corresponding measure of precision [2]. The guidelines encourage the inclusion of information on the direction and size of effect in both absolute and relative measures [2].

Including measures of absolute effects is particularly important for clinical decision making. Expressing treatment effects in relative terms yields apparently larger treatment effects than if absolute terms are used (e.g., a 50% relative risk reduction can mean an absolute risk reduction [ARR] of 1%; i.e., 2−1%), and this difference influences the judgment of clinicians and patients regarding the treatment options [3]. Relative measures may also exaggerate small between-group differences for uncommon events and minimize large differences for common events. Thus, exclusively presenting measures of relative effect without absolute measures can be misleading and present further challenges to patients and clinicians attempting to trade off between desirable and undesirable effects [4−7]. Recent guidelines by the Cochrane Collaboration have also recommended that review authors report estimates of absolute effects in their study abstracts and acknowledged that they are more easily interpretable than relative effects [8].

Although there is a high concordance between abstract and full-text article reporting of single primary outcomes in randomized controlled trials (RCTs) published in top medical journals, little is known about the extent to which abstracts of SRs report the most patient-important outcomes of benefit and harm and the extent to which abstracts report absolute measures of effect [9]. Considering that health care professionals often rely on study abstracts for decision making, we examined the extent to which authors reported absolute effects on patient-important outcomes in the abstracts of Cochrane and non-Cochrane reviews reporting dichotomous outcomes.

2. Methods

2.1. Objectives

1. To examine the extent to which the most patient-important outcomes are reported in the full texts and abstracts of Cochrane and non-Cochrane reviews.
2. As judged from abstracts alone, to examine the use of explicit reporting of patient-important outcomes.
3. As judged from abstracts alone, to examine the extent to which absolute effects are reported for patient-important outcomes.
4. To examine the extent to which the most patient-important outcomes are reported in the full texts and plain language summaries of Cochrane reviews.

2.2. Search strategy

Based on a publically available study protocol [10], which we modified to include a focus on abstract reporting

What is new?

Key findings

- Authors frequently do not report the most patient-important outcomes of benefit and harm in systematic review (SR) abstracts.
- Authors seldom report absolute measures of effect for outcomes of benefit or harm in SR abstracts.

What this adds to what was known?

- Although there is a high concordance between abstract and full-text article reporting of single primary outcomes in randomized trials published in top medical journals, little is known about the extent to which abstracts of SRs report the most patient-important outcomes of benefit and harm and the extent to which abstracts report absolute measures of effect.

What is the implication and what should change now?

- To facilitate health care decision making, journal editors should insist that authors report the most patient-important outcomes of benefit and harm in abstracts using absolute measures of effect.
- Two or more representative absolute effects should be presented if baseline risks vary substantially.

of patient-important outcomes, we searched OVID MEDLINE and the Cochrane Database of Systematic Reviews to identify eligible SRs published in the year 2010. An adapted version of the SR filter designed by the McMaster University Health Information Research Unit was used to retrieve non-Cochrane SRs and Cochrane SRs using the *search by journal* option. Citations were exported into DistillerSR (Evidence Partners, Ottawa, Ontario, Canada).

2.3. Study selection and data extraction

A sample of 2,328 studies was identified by our search. After the removal of duplicates, 2,278 studies were stratified into two groups: Cochrane and non-Cochrane reviews. Within each stratum, repeated random sampling in a 1:1 ratio was used in an iterative process, where pairs of reviewers independently screened studies at the title and abstract and full-text levels from both strata. A sample size of 202 SRs was obtained, of which 12 SRs were further deemed ineligible, leaving 190 eligible SRs. We included studies described as SR or meta-analysis that were published in English; reported a search strategy in at least one database; included a comparison of at least two interventions in humans; reported at least one dichotomous outcome; included at least one RCT; and included only RCTs as primary studies. SRs that only included primary outcomes of harm with no patient-important outcomes of benefit were excluded.

Among eligible SRs, pairs of reviewers selected and extracted data on the population, intervention(s), comparator(s), and the most patient-important outcomes of benefit and harm. Each team of extractors consisted of one more experienced reviewer and one less experienced reviewer. Pretested and standardized forms with extensive instructions were used for eligibility screening, outcome selection, and data abstraction, and any discrepant judgments between reviewers were resolved by consensus, or, if necessary, by a third reviewer. Each reviewer team completed calibration exercises before all screening and abstraction phases to further ensure reliability.

The relative patient importance of outcomes was guided by a hierarchy (Table 1). Mortality (category I), morbidity (category II), and symptoms/quality of life/functional status (category III) were considered patient-important outcomes, whereas surrogate outcomes (category IV) were not. Categories I and II were considered critically important, whereas category III was considered important.

If outcomes of harm or benefit reported by a trial were equally important, the outcome with the largest number of observations was chosen. For composite outcomes, the component of greatest patient importance was selected. When the necessary disaggregated data for a composite outcome of interest was unavailable, we chose the next outcome of greatest patient importance that was reported.

Data abstractors documented the reporting of absolute and relative treatment measures of effect for all eligible review abstracts and if authors used explicit or implicit statements when reporting the most patient-important outcome of benefit and harm. We defined explicit reporting as a direct statement for a specific outcome with an absolute or relative numerical measure. For example, the following statement would be considered explicit reporting: "*In the five studies, 143 people allocated to calcium had a myocardial infarction compared with 111 allocated to placebo (hazard ratio 1.31, 95% confidence interval 1.02 to 1.67, $P = 0.035$).*" We defined implicit reporting as either a broad or an indirect reference to the specific outcome or lacking an absolute or relative numeric measure. For example, the following statement would be considered implicit reporting: "*In the five studies, serious harms were associated with calcium compared to placebo*". As a secondary outcome, we documented the proportion of abstracts and Cochrane plain language summaries that reported the beneficial and harmful outcomes judged as most patient-important outcome from the full text.

Absolute measures of interest included the ARR or risk difference (RD), number needed to treat (NNT), and the proportion of beneficial or harmful events per group. Relative measures of interest included relative risk or rate ratio

Table 1. Hierarchical approach to prioritization of patient-important outcomes

Category	Description	Examples	Critically important or important
I	Mortality	• All-cause mortality • Disease-specific mortality	Critically important
II	Morbidity	• Cardiovascular major morbid events • Other major morbid events (e.g., loss of vision, seizures, fracture, revascularization) • Recurrence/relapse/remission of cancer/disease-free survival • Renal failure requiring dialysis • Hospitalizations • Infections (e.g., pneumonia, gastroenteritis) • Dermatologic (e.g., skin necrosis, Stevens–Johnson syndrome, and toxic epidermal necrolysis) • Rheumatologic disorders such as exacerbation of a rheumatologic disease (e.g., Behcet disease, rheumatoid arthritis)	Critically important
III	Symptoms/quality of life/functional status	• Failure to become pregnant • Successful nursing/breastfeeding • Depression	Important
IV	Surrogate outcomes	• Diagnosis of TB • Viral load • Blood pressure • FEV1 • Physical activity • Postoperative atrial fibrillation • Cognitive function	Somewhat important

Abbreviations: TB, tuberculosis; FEV1, forced expiratory volume in the first second.

(RR), odds ratio, hazard ratio, and relative risk reduction. As some included studies reported both absolute and relative measures in their abstracts, it was noted that the number of measures reported exceeded the number of included studies.

2.4. Data analysis

Frequencies and proportions, and the associated measures of statistical dispersion (interquartile range [IQR]), were calculated for all items, stratified by Cochrane and non-Cochrane SRs. We compared proportions between Cochrane and non-Cochrane SRs with the chi-square test.

We calculated the chance-corrected agreement between reviewers regarding whether an absolute effect estimate was reported for the outcome of interest and interpreted the results according to guidelines proposed by Landis and Koch [11]. Agreement of 0.01–0.20 was considered as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as very good.

All data analyses were performed using SPSS statistical software, version 18.0 (SPSS, Inc., Chicago, IL, USA).

3. Results

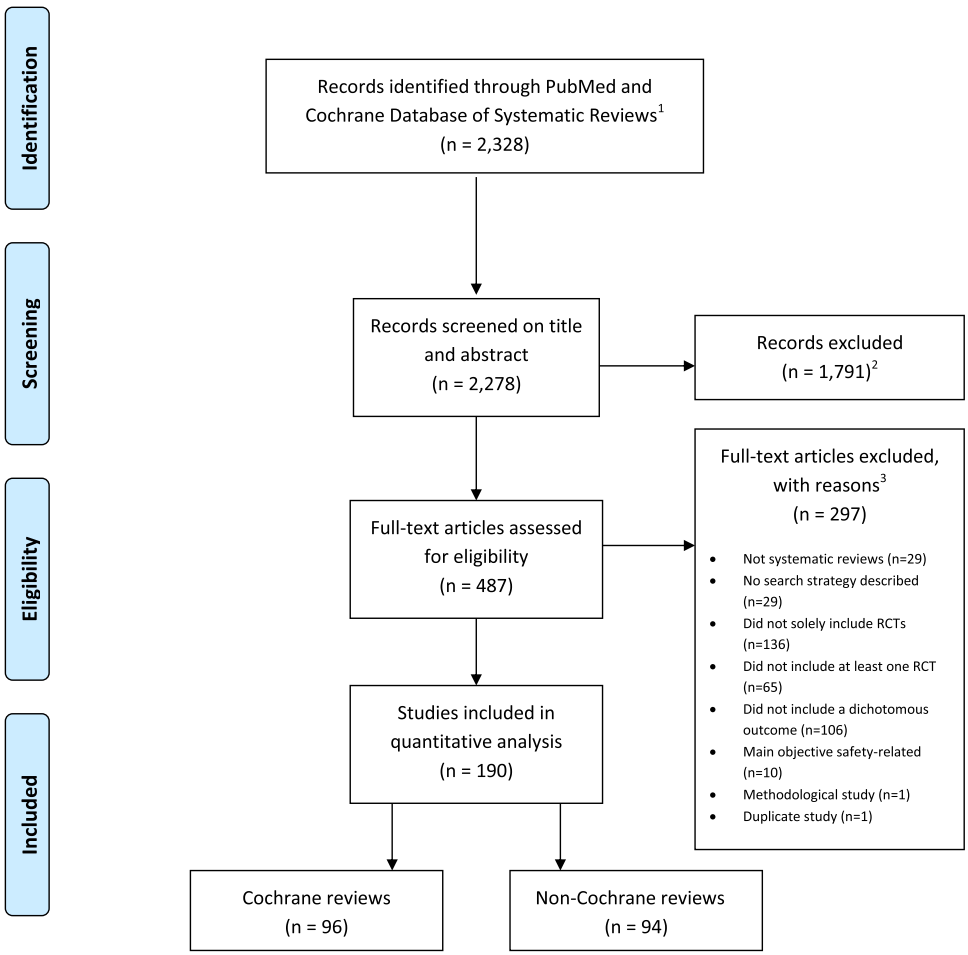
We identified 2,278 potentially eligible articles after removal of duplicates. After stratification between Cochrane and non-Cochrane SRs, repeated random sampling in a 1:1 ratio and initial title and abstract screening, 487 articles were deemed eligible for full-text review, of which 202 proved eligible. Twelve SRs were subsequently excluded, having only reported a harm-related outcome without an eligible outcome of benefit or because of study design, leaving 190 eligible SRs (96 Cochrane and 94 non-Cochrane SRs) (Fig. 1).

The 190 eligible reviews included a median of five RCTs (IQR, 2–9), median number of 851 patients (IQR, 225–2,108), and median of 111.5 events for the outcome of interest (IQR, 31.5–391). Of the 190 reviews, 141 (74.2%) included a focus on a medical area and 55 (29.0%) on a surgical area, with six (3.2%) studies including a focus on both areas. The primary outcome represented a benefit in 172 (90.5%) of the reviews and harm in the remaining 18 (9.5%). In 135 (71.1%) of eligible SRs, the primary outcome chosen by the SR investigators was considered patient important (Tables 2 and 3).

3.1. Reporting of outcomes of benefit in the abstract

Of 190 SRs, 144 (75.8%) reported the outcome of benefit we judged most patient important from the full text in the abstract. Five (2.6%) reported only a composite outcome in the abstract despite reporting disaggregated data in their full text. Thirty-one (16.3%) SRs did not report any of the outcomes of benefit in their abstract. Results were similar between Cochrane and non-Cochrane SRs ($P = 0.91$) (Table 4).

As judged from abstract reporting alone, 161 (84.7%) abstracts reported a patient-important outcome of benefit. Of these, 121 (75.2%) abstracts provided an explicit statement, and 40 (24.8%) provided an implicit statement. Twenty-nine SRs (15.3%) did not report an important outcome of benefit in their abstract. No significant differences were observed between Cochrane and non-Cochrane SRs (Table 4).



[1] Studies were stratified between Cochrane and non-Cochrane groups after the removal of duplicates.

[2] The number excluded includes studies that were screened and proved to be in eligible, and other studies that were not screened after the target sample size of eligible SRs was obtained.

[3] Multiple reasons for exclusion may be selected for a given study; therefore, reasons for exclusions might exceed total number of studies excluded.

Fig. 1. Flow diagram of the literature screening process. RCTS, randomized controlled trials.

One hundred twenty-seven (66.8%) SRs reported a total of 151 absolute and/or relative measures in their abstracts, of which 117 (77.5%) were relative measures and 34 (22.5%) were absolute measures. Eighteen (9.5%) abstracts provided both relative and absolute measures, with ARR/RD combined with RR ($n = 4$) and rate of events per group combined with RR ($n = 6$) being the most common combinations reported. No significant differences were observed between Cochrane and non-Cochrane SRs (Table 4).

3.2. Reporting of outcomes of harm in the abstract

Of 190 SRs, 73 (38.4%) reported critically important outcomes of harm in the full text. Of these, 46 (63.0%) reported this outcome in the abstract. Of our 190 SRs, 97 (51.1%) reported important outcomes of harm in the full text. Of these, 67 (69.1%) reported this outcome in the abstract. No significant differences were observed between Cochrane and non-Cochrane SRs for critically important or important outcomes of harm ($P = 0.13$) (Table 5).

As judged from abstract reporting alone, 87 (45.8%) abstracts reported a patient-important outcome of harm. Of those, an explicit statement was provided in 36 (41.4%) SRs, whereas an implicit statement was provided in 51 (58.6%) SRs. No significant differences were observed between Cochrane and non-Cochrane SRs (Table 5).

Forty-two (22.1%) SRs reported a total of 47 absolute and/or relative measures in their abstracts, of which 41 (87.2%) were relative measures and only 6 (12.8%) were absolute measures. Only two (1.1%) abstracts provided both a

Table 2. General characteristics of the included systematic reviews

Characteristics of reviews	Cochrane *n* = 96 (%)	Non-Cochrane *n* = 94 (%)	Overall *n* = 190 (%)	*P*
Type of analysis developed by authors[a]				
Standard meta-analysis	86 (89.6)	90 (95.7)	176 (92.6)	0.16
Metaregression	3 (3.1)	13 (13.8)	16 (8.4)	**0.009**
Individual patient data meta-analysis	2 (2.1)	4 (4.3)	6 (3.2)	0.44
Network meta-analysis or multiple treatment comparison	1 (1)	0 (0)	1 (0.5)	1
No statistical pooling conducted	6 (6.3)	2 (2.1)	8 (4.2)	0.28
Characteristics of meta-analysis				
Median number of trials included (IQR)	3 (2–8)	7 (4–9)	5 (2–9)	**<0.0001**
Median number of patients included (IQR)	567 (173–1,345)	1,402 (484–3,517)	851 (225–2,108)	**0.003**
Median number of events included (IQR)	59 (16–185)	316 (72–886)	111.5 (31.5–391)	**<0.0001**
Characteristics of meta-analysis[a]				
Medical area	63 (65.6)	78 (83.0)	141 (74.2)	**0.006**
Surgical area	36 (37.5)	19 (20.2)	55 (29.0)	**0.009**
Intervention for comparison of interest				
Pharmacologic	68 (70.8)	53 (56.4)	121 (63.7)	**0.04**
Nonpharmacologic	28 (29.2)	41 (43.6)	69 (36.3)	
Control for comparison of interest				
Pharmacologic	35 (36.4)	28 (29.8)	63 (33.2)	0.61
Nonpharmacologic	33 (34.4)	37 (39.4)	70 (36.8)	
Placebo	28 (29.2)	29 (30.8)	57 (30.0)	
Use of GRADE				
Yes	27 (28.1)	3 (3.2)	30 (15.8)	**<0.0001**
No	68 (70.8)	90 (95.7)	158 (83.1)	
Unclear	1 (1.1)	1 (1.1)	2 (1.1)	
Evaluation of risk of bias				
Risk of bias tool	82 (85.4)	9 (9.6)	91 (47.9)	**<0.0001**
By dimensions	8 (8.3)	29 (30.9)	37 (19.5)	
Jadad or other scales	4 (4.2)	29 (30.9)	33 (17.4)	
Not evaluated	0 (0.0)	16 (17.0)	16 (8.4)	
Other	2 (2.1)	11 (11.6)	13 (6.8)	
Funding				
For profit	2 (2.1)	4 (4.3)	6 (3.2)	0.44
Not for profit	62 (64.6)	33 (35.1)	95 (50)	**<0.0001**
Not funded	10 (10.4)	11 (11.7)	21 (11.1)	0.78
Not reported	23 (24)	47 (50)	70 (36.8)	**0.0002**
Did authors report ties to industry?				
Yes	20 (20.8)	15 (16)	35 (18.4)	**0.0002**
No	57 (59.4)	34 (36.2)	91 (47.9)	
Not reported	19 (19.8)	43 (45.7)	62 (32.6)	
Unclear	0 (0)	2 (2.1)	2 (1.1)	
Mean quality score on the AMSTAR instrument (SD)	10 (9–10)	7 (6–8)	9 (7–10)	**<0.0001**

Abbreviations: IQR, interquartile range; GRADE, Grades of Recommendation, Assessment, Development and Evaluation; AMSTAR, A Measurement Tool to Assess Systematic Reviews; SD, standard deviation.

Statistically significant *P*-values are indicated in bold.

[a] The questions allow multiple selection; therefore, sum of percentages might exceed 100%.

relative measure (odds ratio and RR) and an absolute measure (ARR/RD). No significant differences were observed between Cochrane and non-Cochrane SRs (Table 5).

3.3. Reporting of outcomes of benefit and harm in Cochrane plain language summaries

Among 94 (97.9%) Cochrane SRs reporting a patient-important outcome of benefit in their full text, 43 (45.7%) explicitly reported the most patient-important outcome of benefit from the full text in their plain language summary, 28 (29.8%) reported the outcome as an implicit statement, and 23 (24.5%) did not report the outcome in the summary (Table 6).

Among 68 (70.8%) Cochrane SRs reporting a patient-important outcome of harm in their full text, 20 (29.4%) explicitly reported the most patient-important outcome of harm from the full text in their plain language summary, 17 (25.0%) reported the outcome as an implicit statement, and 31 (45.6%) did not report the outcome in the summary (Table 6).

4. Discussion

4.1. Principal findings

Among 190 eligible SRs (96 Cochrane and 94 non-Cochrane) reporting dichotomous outcomes, we found that

Table 3. Patient-important outcomes in the included systematic reviews

Characteristics of reviews	Cochrane $n = 96$ (%)	Non-Cochrane $n = 94$ (%)	Overall $n = 190$ (%)	*P*
Is the most patient-important outcome the primary outcome for the authors?				
Yes	64 (66.7)	45 (47.9)	109 (57.4)	**<0.0001**
No, but some other outcome specified as primary	21 (21.9)	12 (12.8)	33 (17.4)	
Authors did not specify a primary outcome	11 (11.5)	37 (39.3)	48 (25.3)	
Is the primary outcome of the authors patient important?				
Yes	81 (84.4)	54 (57.5)	135 (71.1)	**<0.0001**
No	6 (6.3)	5 (5.3)	11 (5.8)	
Authors did not specify a primary outcome	9 (9.5)	35 (36.8)	44 (23.2)	
Most patient-important outcome one of benefit or harm?				
Beneficial outcome	85 (88.5)	87 (92.6)	172 (90.5)	0.35
Harm outcome	11 (11.6)	7 (7.4)	18 (9.5)	
Relative effect estimates for the most patient-important outcome				
Relative risk	58 (60.4)	53 (56.4)	111 (58.4)	0.57
Odds ratio	24 (25)	28 (29.8)	52 (27.4)	0.46
Hazard ratio	8 (8.3)	3 (3.2)	11 (5.8)	0.21
RRR	0 (0)	1 (1.1)	1 (0.5)	1
Rate ratio	1 (1)	1 (1.1)	2 (1.1)	1
Other	5 (5.2)	9 (9.6)	14 (7.4)	0.25

Abbreviation: RRR, relative risk reduction.
Statistically significant *P*-values are indicated in bold.

authors frequently did not report the most patient-important outcomes of benefit and harm in their abstracts and plain language summaries. Authors seldom reported absolute measures of effect for either beneficial outcomes or harmful outcomes. These findings proved similar in Cochrane and non-Cochrane SRs.

4.2. Previous findings

The findings of our study are consistent with a structured review of general medical and public health journals that examined the frequency of absolute and relative measure reporting in health inequalities research [12]. Evaluating 344 articles published in 2009 among the top four general medicine and top four public health journals, the review found that only 138 reported numeric data in the abstract; among these, 88% reported a relative measure of effect, 9% reported only an absolute measure, and 2% reported both [12]. Similarly, an evaluation of 182 SR abstracts published in 2009, Beller et al. [9] found that only 137 reported numeric data in the abstract; among these, 45 (25%) did not report an absolute measure or a relative measure, and 87 (48%) reported only relative measures. In contrast, 28 (15%) abstracts reported only absolute measures, 14 (8%) reported data allowing both absolute and relative measures to be calculated, and only 8 (4%) reported both measures [9].

Our study builds on the findings of these two studies but focuses on patient-important outcome reporting in

Table 4. Reporting of patient-important outcomes of benefit

Variable of interest	Cochrane	Non-Cochrane	Overall	*P*
Reporting of the patient-important outcome of benefit judged most important in the abstract	$n = 96$ (%)	$n = 94$ (%)	$n = 190$ (%)	
Most important outcome in the full text is also reported in the abstract	73 (76.0)	71 (75.5)	144 (75.8)	0.91
Disaggregated data reported in the full text, but only a composite reported in the abstract	3 (3.1)	2 (2.1)	5 (2.6)	
Reported in the full text but not in the abstract	16 (16.7)	15 (16.0)	31 (16.3)	
Article has no patient-important outcomes reported in the full text	4 (4.2)	6 (6.4)	10 (5.3)	
Explicit vs. implicit reporting of the outcome of benefit in the abstract	$n = 96$ (%)	$n = 94$ (%)	$n = 190$ (%)	
Explicit reporting: direct statement, with numerical value provided	63 (65.6)	58 (61.7)	121 (63.7)	0.76
Implicit reporting: direct statement, with no numerical value provided	13 (13.5)	10 (10.6)	23 (12.1)	0.50
Implicit reporting: indirect statement, with numerical value provided	1 (1.0)	5 (5.3)	6 (3.2)	0.21
Implicit reporting: indirect statement, with no numerical value provided	8 (8.3)	3 (3.2)	11 (5.8)	0.21
No patient-important beneficial outcomes reported in the abstract	11 (11.5)	18 (19.2)	29 (15.3)	0.23
Absolute vs. relative reporting of the outcome of benefit in the abstract	$n = 96$ (%)	$n = 94$ (%)	$n = 190$ (%)	
Relative measure provided in the abstract (RR, OR, HR, RRR, or rate ratio)	57 (53.3)	60 (56.1)	117 (54.7)	0.38
Absolute measure provided in the abstract (ARR/RD, NNT/NNTB, NNH/NNTH, rate of events per group, or rate difference)	18 (16.8)	16 (15.0)	34 (15.9)	0.82
No patient-important beneficial outcomes reported in the abstract	11 (10.3)	18 (16.8)	29 (13.6)	0.23
No numerical value provided	21 (19.6)	13 (12.1)	34 (15.9)	0.13

Abbreviations: RR, relative risk; OR, odds ratio; HR, hazard ratio; RRR, relative risk reduction; ARR, absolute risk reduction; RD, risk difference; NNT, number needed to treat; NNTB, number needed to treat (benefit); NNH, number needed to harm; NNTH, number needed to treat (harm).

Table 5. Reporting of patient-important outcomes of harm

Variable of interest	Cochrane	Non-Cochrane	Overall	*P*
Patient "critically important" outcome of harm reporting in the abstract	*n* = 96 (%)	*n* = 94 (%)	*n* = 190 (%)	
"Critically important" outcomes of harm are explicitly reported in the abstract	18 (18.8)	15 (16.0)	33 (17.4)	0.13
"Critically important" outcomes of harm are implicitly reported in the abstract	8 (8.3)	5 (5.3)	13 (6.8)	
"Critically important" outcomes of harm are not reported in the abstract	18 (18.8)	9 (9.6)	27 (14.2)	
No "critically important" outcomes of harm are reported in the full text	52 (54.2)	65 (69.1)	117 (61.6)	
Patient-important outcome of harm reporting in the abstract	*n* = 96 (%)	*n* = 94 (%)	*n* = 190 (%)	
"Important" outcomes of harm are explicitly reported in the abstract	25 (26.0)	15 (16.0)	40 (21.1)	0.09
"Important" outcomes of harm are implicitly reported in the abstract	14 (14.6)	13 (13.8)	27 (14.2)	
"Important" outcomes of harm are not reported in the abstract	19 (19.8)	11 (11.7)	30 (15.8)	
No "important" outcomes of harm are reported in the full text	38 (39.6)	55 (58.5)	93 (48.9)	
Explicit vs. implicit reporting of the most patient-important outcome of harm in the abstract	*n* = 96 (%)	*n* = 94 (%)	*n* = 190 (%)	
Explicit reporting: direct statement, with numerical value provided	18 (18.8)	18 (19.1)	36 (18.9)	1
Implicit reporting: direct statement, with no numerical value provided	14 (14.6)	10 (10.6)	24 (12.6)	0.38
Implicit reporting: indirect statement, with numerical value provided	2 (2.1)	4 (4.3)	6 (3.2)	0.68
Implicit reporting: indirect statement, with no numerical value provided	15 (15.6)	6 (6.4)	21 (11.1)	0.04
No important or critically important harmful events reported in the abstract	47 (49.0)	56 (59.6)	103 (54.2)	0.11
Absolute vs. relative reporting of the most patient-important outcome of harm in the abstract	*n* = 96 (%)	*n* = 94 (%)	*n* = 190 (%)	
Relative measure provided in abstract (RR, OR, HR, RRR, or rate ratio)	21 (20.8)	20 (21.3)	41 (21.0)	0.72
Absolute measure provided in the abstract (ARR/RD, NNT/NNTB, NNH/NNTH, rate of events per group, or rate difference)	4 (4.0)	2 (2.1)	6 (3.1)	0.62
No important or critically important harmful events reported in the abstract	47 (46.5)	56 (59.6)	103 (52.8)	0.11
No numerical value reported in the abstract	29 (28.7)	16 (17.0)	45 (23.1)	0.04

Abbreviations: RR, relative risk; OR, odds ratio; HR, hazard ratio; RRR, relative risk reduction; ARR, absolute risk reduction; RD, risk difference; NNT, number needed to treat; NNTB, number needed to treat (benefit); NNH, number needed to harm; NNTH, number needed to treat (harm).

Cochrane and non-Cochrane SRs published in 2010 across all medical and surgical fields. In line with their findings, we also found an overreliance on relative vs. absolute measures in SR abstracts, and only a small number of SR abstracts that reported both types of measures.

4.3. Strengths and limitations

Our study has several strengths. We established explicit eligibility criteria, conducted sensitive searches to identify relevant data, and conducted eligibility screening and data abstraction independently and in duplicate with discussion to resolve discrepancies and establish consensus. We chose a large and representative sample of both Cochrane and non-Cochrane SRs, making the findings of our study broadly applicable. Unlike previous studies in the area [12], our findings are not restricted to SRs from top journals. We assessed the reporting and presentation format of patient-important outcomes in abstracts, evaluating the use of absolute and relative measures and explicit and implicit statements, issues that have not been previously studied.

Table 6. Reporting of most patient-important outcomes in Cochrane plain language summaries

Variable of interest	*n* = 96 (%)
Most patient-important outcome of benefit reporting in the plain language summary	
Yes, reported as explicit statement	43 (44.8)
Yes, reported as implicit statement	28 (29.2)
Patient-important outcomes reported in the full text but not in the plain language summary	23 (24.0)
Article has no patient-important outcomes reported in the full text	2 (2.1)
Most patient-important outcome of harm reporting in the plain language summary	
Yes, reported as an explicit statement	20 (20.8)
Yes, reported as an implicit statement	17 (17.7)
Patient-important harmful event reported in the full text but not in the plain language summary	31 (32.3)
Article has no patient-important harmful event reported in the full text	28 (29.2)

Our study has several limitations. First, there was considerable judgment involved in many of the data-abstracting decisions, in particular the judgment of outcomes of most importance. We did, however, provide extensive instructions and conducted calibration exercises.

Second, our search was limited to MEDLINE and the Cochrane Database of Systematic Reviews, which might limit the generalizability of our findings. It is, however, likely that non–MEDLINE-indexed journals do no better than the MEDLINE-indexed journals we sampled.

Third, it may be that the quality of reporting has improved in recent years as a result of guidance from Grades of Recommendation, Assessment, Development and Evaluation (GRADE) and the Methodological Expectations of Cochrane Intervention Reviews project, which encourage reporting of effect estimates in both relative and absolute effects. Furthermore, the increased uptake of GRADE methodology and associated Summary of Findings table in SRs, particularly Cochrane SRs, is likely to have increased the use of absolute measures in SRs, and by extension, in their abstracts.

A 2012 study involving a cohort of Cochrane SRs showed that 45% of the SRs used the GRADE approach

to interpret the overall quality of a body of evidence vs. only 6% of non-Cochrane SRs [13]. Considering the inclusion of GRADE as a surrogate of absolute estimates in the SR, we may estimate an approximate increase of 6% in absolute measure reporting in the SR over 2 years since our 2012 data. Given that it is very unlikely that non-Cochrane SRs have shown greater improvement than Cochrane SRs in the interval, it is unlikely that our study markedly underestimates reporting of absolute measures in the abstract.

4.4. Implications and future directions

Authors should, however, be careful in communicating the size of treatment effects. Assuming results, in relative terms, are significant: to calculate RDs, authors should ideally apply the relative estimate (e.g., RR) to the baseline risk from well-designed, large, and observational studies. If such observational studies are not available, authors should consider the variation in the baseline risk among included studies in the SR. If there is little baseline variation, authors can use the median control group risk from the included studies. If large variation exists, authors can consider using two or more representative baseline risks from the included studies [14].

For example, a meta-analysis of probiotics for the prevention of *Clostridium difficile* infection included 21 RCTs, estimated an RR of 0.39, and observed baseline risks ranging from 0% to 40%, with a mean risk in control groups of 3.8% [15,16]. Using the mean baseline risk of 3.8%, the RR of 0.39 corresponds to an ARR of 2.3% and an NNT of 43. Using the same RR of 0.39, based on the control group of the largest RCT to date (n = 1,488), the baseline risk of 1.2% corresponds to an ARR of 0.73% and an NNT of 137 [17]. In a second example drawn from an actual abstract: "patients randomized to primary percutaneous coronary intervention (PPCI) had lower mortality than did patients randomized to fibrinolysis (5.3% vs 7.9%, adjusted odds ratio 0.63, 95% CI 0.42-0.84, $P < 0.001$). The interaction between risk score and allocated treatment interaction term had no significant contribution ($P = 0.52$) to the model, indicating that the relative mortality reduction by PPCI was similar at all levels of estimated risk. In contrast, the absolute risk reduction was strongly related to estimated risk at baseline: the numbers needed to treat to prevent a death by PPCI vs. fibrinolysis was 516 in the lowest quartile of estimated risk compared with only 17 in the highest quartile" [18].

As illustrated previously, the absolute size of the treatment effect can vary substantially depending on the baseline risk. In cases like the two described, reporting only the pooled mean ARR can be as uninformative or misleading as not reporting ARR at all, indicating that the assessment of the baseline risk in clinical decision making is crucial [19,20].

Our findings suggest that major improvements in the reporting of absolute measures for patient-important outcomes in Cochrane and non-Cochrane SR abstracts are necessary. The PRISMA consensus statement on recommended items to include in the reporting of SRs and meta-analyses aims to promote transparency and complete reporting [21]. In its elaboration document, PRISMA encourages authors to report main results in the data synthesis section of their abstract, using numerical results with confidence intervals for the SRs' most important outcomes, and ideally, with specifications of the number of studies and participants included in the analyses [22]. This is reaffirmed in the PRISMA for abstracts checklist published in 2013 [2], which suggests that in addition to presenting the baseline risk, the presentation of results for the main outcomes of benefit and harm should include the number of studies and participants, summary measures, confidence intervals, direction of effect and size of effect (e.g., lower, fewer, reduced; and greater, more, increased). With respect to absolute measures of effect, the PRISMA statement provides guidance (although somewhat limited): when a percentage is used, the baseline risk should also be shown, which allows the reader to see what the absolute benefit or harm is, and calculate whichever measures they choose. The guideline also suggests that results be presented in terms meaningful to clinicians and patients, including kilograms, days, and percentages [2].

PRISMA for abstracts neither, however, does provide guidance to authors regarding the inclusion of the most patient-important outcomes of benefit and harm nor does it emphasize the need to report absolute measures of effect. The low proportions of SRs reporting patient-important outcomes and absolute measures are perhaps unsurprising in the light of the guidance provided to authors of SRs.

5. Conclusions

Strong and explicit guidance in authoritative reporting guidelines, such as those provided as part of the PRISMA abstract statement, should better address patient-important outcome reporting with absolute measures of effect. To facilitate health care decision making, journal editors should insist, in abstracts, that authors report all prespecified outcomes, which should include the most patient-important outcomes of benefit and harm in abstracts using absolute measures of effect. When interpreting research results based on abstract reporting, users of SRs should be aware of the merit of absolute effects for decision making.

Acknowledgments

Pablo Alonso-Coello is funded by a Miguel Servet research contract from the Instituto de Salud Carlos III (CP09/00137); Laura Martínez García and Andrea Juliana Sanabria are funded by a Río Hortega research contract from the Instituto de Salud Carlos III (CM10/00014 and CM12/00168); Matthias Briel is supported by santésuisse and the Gottfried and Julia Bangerter-Rhyner Foundation; Jason Busse is supported by a New Investigator Award from the Canadian Institutes of Health Research and

Canadian Chiropractic Research Foundation; and the work of Kari Tikkinen was supported by the Academy of Finland (276046), Competitive Research Funding of the Helsinki and Uusimaa Hospital District, Finnish Cultural Foundation, Finnish Medical Foundation, Jane and Aatos Erkko Foundation, and Sigrid Jusélius Foundation.

References

[1] The PLoS Medicine Editors. The impact of open access upon public health. PLoS Med 2006;3:e252.

[2] Beller EM, Glasziou PP, Altman DG, Hopewell S, Bastian H, Chalmers I, et al. PRISMA for abstracts: reporting systematic reviews in journal and conference abstracts. PLoS Med 2013;10:e1001419.

[3] Barratt A, Wyer PC, Hatala R, McGinn T, Dans AL, Keitz S, et al. Tips for learners of evidence-based medicine: 1. Relative risk reduction, absolute risk reduction and number needed to treat. CMAJ 2004;171:353−8.

[4] Replogle WH, Johnson WD. Interpretation of absolute measures of disease risk in comparative research. Fam Med 2007;39:432−5.

[5] Citrome L. Quantifying clinical relevance. Innov Clin Neurosci 2014;11:26−30.

[6] Citrome L. Relative vs. absolute measures of benefit and risk: what's the difference? Acta Psychiatr Scand 2010;121:94−102.

[7] Akobeng AK. Understanding measures of treatment effect in clinical trials. Arch Dis Child 2005;90:54−6.

[8] Methodological Expectations of Cochrane Intervention Reviews (MECIR). Standards for the reporting of new Cochrane Intervention Reviews. Report. v.1.1, 2012.

[9] Beller EM, Glasziou PP, Hopewell S, Altman DG. Reporting of effect direction and size in abstracts of systematic reviews. JAMA 2011;306:1981−2.

[10] Alonso-Coello P, Carrasco-Labra A, Brignardello-Petersen R, Neumann I, Akl EA, Sun X, et al. A methodological survey of the analysis, reporting and interpretation of Absolute Risk ReductiOn in systematic revieWs (ARROW): a study protocol. Syst Rev 2013;2:113.

[11] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159−74.

[12] King NB, Harper S, Young ME. Use of relative and absolute effect measures in reporting health inequalities: structured review. BMJ 2012;345:e5774.

[13] Hopewell S, Boutron I, Altman DG, Ravaud P. Incorporation of assessments of risk of bias of primary studies in systematic reviews of randomised trials: a cross-sectional study. BMJ Open 2013;3:e003342.

[14] Alonso-Coello P, Carrasco-Labra A, Brignardello-Petersen R, Neumann I, Akl EA, Vernooij RW, et al. Systematic reviews suffer from major limitations in reporting absolute effects. J Clin Epidemiol 2016;72:16−26.

[15] Johnston BC, Ma SS, Goldenberg JZ, Thorlund K, Vandvik PO, Loeb M, et al. Probiotics for the prevention of Clostridium difficile-associated diarrhea: a systematic review and meta-analysis. Ann Intern Med 2012;157:878−88.

[16] Daneman N. A probiotic trial: tipping the balance of evidence? Lancet 2013;382:1228−30.

[17] Allen SJ, Wareham K, Wang D, Bradley C, Hutchings H, Harris W, et al. Lactobacilli and bifidobacteria in the prevention of antibiotic-associated diarrhoea and Clostridium difficile diarrhea in older inpatients (PLACIDE): a randomised, double-blind, placebo-controlled, multicentre trial. Lancet 2013;382:1249−57.

[18] de Boer SP, Barnes EH, Westerhout CM, Simes RJ, Granger CB, Kastrati A, et al. High-risk patients with ST-elevation myocardial infarction derive greatest absolute benefit from primary percutaneous coronary intervention: results from the Primary Coronary Angioplasty Trialist versus thrombolysis (PCAT)-2 collaboration. Am Heart J 2011;161:500−7.

[19] Iorio A, Spencer FA, Falavigna M, Alba C, Lang E, Burnand B, et al. Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. BMJ 2015;350:h870.

[20] Spencer FA, Iorio A, You J, Murad MH, Schünemann HJ, Vandvik PO, et al. Uncertainties in baseline risk estimates and confidence in treatment effects. BMJ 2012;345:e7401.

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

[22] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700.