Are outcomes reported in surgical randomized trials patient-important? A systematic review and meta-analysis

Sam Adie, MBBS, MSpMed, MPH, PhD
Ian A. Harris, MBBS, MMed, PhD
Justine M. Naylor, PhD
Rajat Mittal, BSc(Med), MBBS, MMed(Clin Epi) MS

See also the commentary by Vinden on page 81.

Accepted Sept. 21, 2016; Early-released Feb. 1, 2017

Correspondence to:
S. Adie
Orthopaedic Department
Liverpool Hospital
Locked Bag 7103, Liverpool BC
NSW Australia 1871
sam.adie@gmail.com

DOI: 10.1503/cjs.010616

Background: The dangers of using surrogate outcomes are well documented. They may have little or no association with their patient-important correlates, leading to the approval and use of interventions that lack efficacy. We sought to assess whether primary outcomes in surgical randomized controlled trials (RCTs) are more likely to be patient-important outcomes than surrogate or laboratory-based outcomes.

Methods: We reviewed RCTs assessing an operative intervention published in 2008 and 2009 and indexed in MEDLINE, EMBASE or the Cochrane Central Register of Controlled Trials. After a pilot of the selection criteria, 1 reviewer selected trials and another reviewer checked the selection. We extracted information on outcome characteristics (patient-important, surrogate, or laboratory-based outcome) and whether they were primary or secondary outcomes. We calculated odds ratios (OR) and pooled in random-effects meta-analysis to obtain an overall estimate of the association between patient importance and primary outcome specification.

Results: In 350 included RCTs, a total of 8258 outcomes were reported (median 18 per trial. The mean proportion (per trial) of patient-important outcomes was 60%, and 66% of trials specified a patient-important primary outcome. The most commonly reported patient-important primary outcomes were morbid events (41%), intervention outcomes (11%), function (11%) and pain (9%). Surrogate and laboratory-based primary outcomes were reported in 33% and 8% of trials, respectively. Patient-important outcomes were not associated with primary outcome status (OR 0.82, 95% confidence interval 0.63–1.1, $I^2 = 21\%$).

Conclusion: A substantial proportion of surgical RCTs specify primary outcomes that are not patient-important. Authors, journals and trial funders should insist that patient-important outcomes are the focus of study.

Contexte : Les dangers de l'utilisation de critères de substitution sont bien documentés. Ils peuvent avoir peu de liens, voire aucun, avec leurs corrélats importants pour le patient, menant à l'approbation et à l'utilisation d'interventions inefficaces. Nous avons tenté de déterminer si les résultats primaires d'essais cliniques randomisés en chirurgie sont plus susceptibles d'être des résultats importants pour le patient que des critères de substitution ou des résultats de laboratoire.

Méthodes : Nous avons examiné des essais cliniques randomisés portant sur l'évaluation d'une intervention chirurgicale, publiés en 2008 et 2009 et répertoriés dans MEDLINE, EMBASE ou le Cochrane Central Register of Controlled Trials. Après l'essai du critère de sélection, un examinateur a choisi les essais et un autre examinateur a vérifié la sélection. Nous avons obtenu les renseignements sur les caractéristiques des résultats (importants pour le patient, de substitution ou de laboratoire) et déterminé s'il s'agissait de résultats primaires ou secondaires. Nous avons calculé le rapport des cotes (RC) et regroupé une méta-analyse à effets aléatoires afin d'obtenir une estimation globale du lien entre l'importance pour le patient et la spécification du résultat primaire.

Résultats : Un total de 8258 résultats ont été signalés dans les 350 essais cliniques randomisés inclus (pour une médiane de 18 par essai). La proportion moyenne (par essai) de résultats importants pour le patient était de 60 %, et 66 % des essais précisaient un résultat primaire important pour le patient. Les résultats primaires importants pour le patient les plus couramment signalés étaient les événements morbides (41 %), les résultats liés à une intervention (11 %), le fonctionnement (11 %) et la douleur (9 %). Des résultats primaires de substitution ou de laboratoire ont été signalés dans 33 % et 8 % des essais, respectivement. Les résultats importants pour le patient n'étaient pas associés à la situation du résultat primaire (RC 0,82, intervalle de confiance de 95 %, 0,63–1,1, $I^2 = 21$ %).

Conclusion : Un nombre important d'essais cliniques randomisés en chirurgie précisent des résultats primaires qui ne sont pas importants pour le patient. Les auteurs, les revues et les organismes de financement des essais devraient insister pour que les résultats importants pour le patient soient l'objet principal de l'étude.

Ideally, outcome measurements in clinical trials are important to patients, and therefore immediately relevant to clinical practice. Outcome measurements must be valid and reliable, which poses some difficulty when measuring patient-centred outcomes, such as function or health-related quality of life (HRQoL).[1,2] Other patient-important outcomes, such as death and morbidity may require trials with large sample sizes and long-term follow-up.[2,3] Thus researchers may choose surrogate outcomes that are easier to measure, but only reflect an issue important to the patient, and are not necessarily important themselves.

Surrogate outcomes are defined as "a laboratory measurement or a physical sign used as a substitute for a clinically meaningful end point that measures directly how a patient feels, functions or survives."[3] In diabetes trials, only 18% of primary outcomes were patient-important,[4] and other specialties had similarly low rates after empirical reviews.[5,6] The dangers of using surrogate outcomes are well documented. They may have little or no association with their patient-important correlates, leading to the approval and use of interventions that lack efficacy.[7,8] Of greater concern are approved interventions that are in fact harmful, and the use of surrogate outcomes has been blamed for unnecessary deaths.[9–11]

Little is known about the patterns of outcome reporting in surgical trials. We performed a systematic review and meta-analysis of published randomized controlled trials (RCTs) of surgical interventions. Our primary aim was to determine the proportion of primary outcomes in surgical trials that were patient-important.

Methods

Design and study selection

The methods for this systematic review and meta-analysis were prespecified in a protocol as part of a doctoral thesis. This study is reported here according to the PRISMA statement guidelines.[12] We obtained ethics approval for the present study from the South Western Sydney Local Health District, Human Research Ethics Committee.

We included RCTs published in English and in full text, conducted on humans (not cadavers), that compared a surgical intervention to any other intervention. We defined a surgical intervention as any procedure that requires surgical training and is usually performed by a surgeon of any subspecialty recognized by the Royal Australasian College of Surgeons. This included upper and lower gastrointestinal, transplant, cardiothoracic, neuro-, ear nose and throat, pediatric, plastic and reconstructive, urological, vascular and orthopaedic surgery. Obstetric/gynecologic, ophthalmic and dental surgeries were excluded. Injections of any material, applications of splints, and interventions purely for diagnostic purposes were also excluded.

We devised an electronic search strategy with the help of a medical librarian associated with the Cochrane Collaboration. A randomized trial filter based on the Cochrane highly sensitive search strategy[13,14] was combined with a surgical intervention filter (Appendix 1, available at canjsurg.ca). We searched MEDLINE via Ovid (2005–week 3, May 2009), EMBASE via Ovid (2005–week 21, 2009) and CENTRAL via Wiley Interscience (2005–Issue 2, 2009).

Study identification began with the most recent reference and proceeded backward in time. We aimed to include the 350 most recently published surgical RCTs. Using 1000 references, study identification was piloted by 2 authors (S.A., I.A.H.) in order to resolve issues with interpretation of the eligibility criteria. We screened the titles and abstracts of references and retrieved the full text of potentially eligible articles. Studies were included only after assessment of their full text. The pilot search resulted in almost perfect agreement between the 2 assessors (n = 1000, κ = 0.85, 95% confidence interval [CI] 0.77–0.93), and thereafter study identification was performed by 1 author (S.A.), in an identical process.

Data extraction

We created an electronic proforma for data extraction, after piloting by 3 authors (S.A., R.M., J.N.) using a sample

Box 1: Operational definitions of patient-important outcomes

Patient-important outcomes

Outcomes that are likely to be informative to patients and clinicians and measure factors directly related to patient health. Includes the following:

- Mortality/survival (e.g., 30-d mortality, 5-yr survival)
- Pain (e.g., visual analogue score pain; incidence of a painful symptom, such as dysuria or headache; questionnaire resulting in an aggregate score of pain, such as the Western Ontario and McMaster Universities Arthritis Index pain subscale)
- Function (e.g., validated measures of function, such as the International Index of Erectile Function, or the New York Heart Association Functional Classification)
- Quality of life (e.g., validated measures, such as the Short Form 36 survey, or the EuroQol 5 dimension survey)
- Any morbid event or symptom (e.g., incidence of wound infection, fracture nonunion, incontinence), or a measure of their opposites (e.g., wound healing, fracture union, continence)
- Patient satisfaction (e.g., a survey of overall patient satisfaction with their surgical procedure, or satisfaction with cosmesis)
- Any intervention to address the previous 6 outcomes (e.g., use of analgesia for pain, catheterization for urinary retention, interventions to restore vascular patency, revision surgery)

Surrogate outcomes

Outcomes that may indicate progression or an increased risk of a patient-important outcome, but are not intrinsically important to patients (e.g., operative duration for any procedure, urine flow rate for patients with prostatism, hemodynamic measurements after coronary bypass, fracture alignment after operative fixation, surgeon-reported "success" of a procedure)

Laboratory outcomes

Nonclinical tests that measure physiologic parameters without any direct or tangible effects on patients (e.g., inflammatory markers after surgery, troponin after coronary bypass, cholesterol levels after obesity surgery, amylase/lipase after pancreatic surgery)

of 10 trials. After calibration of the data, 1 author (S.A.) extracted the data. Another researcher (R.M.) checked a random sample of 100 included RCTs, and disagreements were resolved by discussion. One data point from 5 trials (< 1% of the checked sample) was changed after double-checking, and no data relating to classification of patient importance were changed.

Data items

We extracted trial-level characteristics, including author background in epidemiology/statistics, study type (superiority/noninferiority), study design (parallel/split body/crossover/factorial), journal impact factor, total sample size, multicentre status and trial registration. Risk of bias domains were also extracted, including adequacy of random sequence generation, allocation concealment, blinding, reporting of attrition and source of funding. Operational definitions of these variables may be found in Appendix 1.

We defined an outcome as a variable used to compare the randomized groups in a trial in order to assess the efficacy or harm of an intervention.[15] We extracted all outcomes in each trial and recorded outcome-level characteristics, including patient importance and whether the outcome was specified as primary or secondary.

Individual trial outcomes were classified as patient-important, surrogate, or laboratory measurements.[4] Box 1 presents the operational definitions used, with examples from various surgical specialties. Patient-important outcomes included measurements of mortality/survival, pain, function, HRQoL, any morbid event, patient satisfaction and any intervention used to address these.[4] Surrogate outcomes did not meet the criteria for being patient-important, but instead were indicators of progression or an increased risk of a patient-important outcome.[3] Laboratory outcomes were

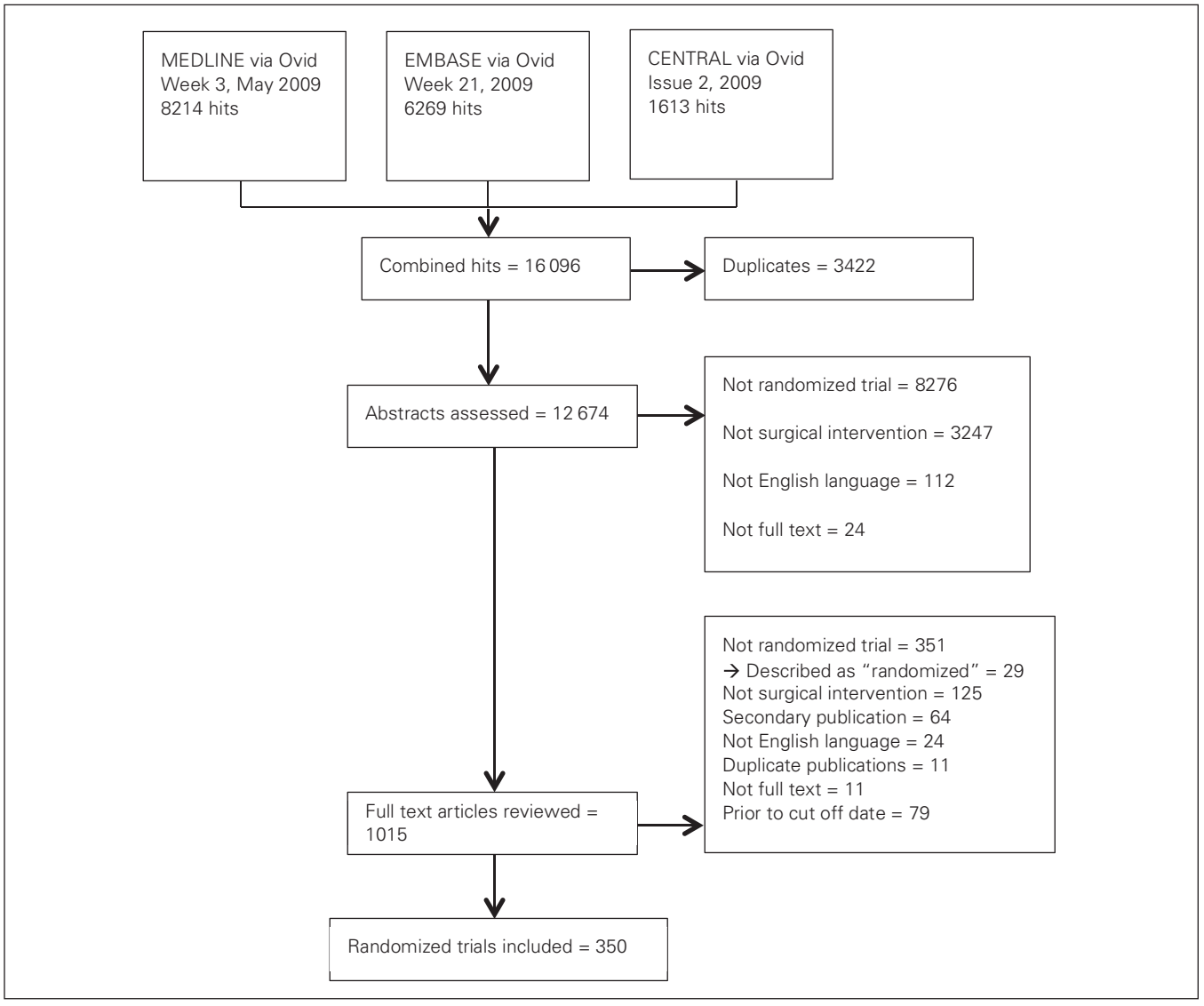

Fig. 1. Selection of studies for inclusion in the review and meta-analysis.

defined as nonclinical tests that measure physiologic parameters without any direct or tangible effects on patients. When composite outcomes were reported, the components of that outcome were graded separately according to the previously mentioned criteria.[4]

Each outcome was also recorded as to whether it was specified as primary, secondary, or unclear. A primary outcome was either used in a sample size calculation, defined explicitly as such in the text (using the word "primary" or a synonym), or was stated explicitly in an aims/hypothesis statement. When a primary outcome was specified, we regarded other outcomes as secondary outcomes. When no primary or secondary outcomes were specified in a trial, we marked that trial's outcomes as unclear.

Statistical analysis

We recorded the total number of outcomes per trial, and calculated the proportion of patient-important outcomes (as well as a composite of any patient-important outcome), surrogate outcomes and laboratory outcomes at the trial level. Mean proportions were calculated for the whole sample of trials.

For each trial, we populated a contingency (2×2) table with that trial's outcomes, characterizing whether each outcome was patient-important or not, and whether each outcome was specified as primary or secondary. Trials that did not specify primary and/or secondary outcomes were not eligible for this analysis. If the contingency table contained a single zero cell or 2 diagonal zero cells, we added 0.5 to all 4 cells as per standard methods found in statistical packages.[16] When a whole row or column contained zero cells, an odds ratio (OR) was incalculable and that trial was excluded from this statistical analysis. An OR greater than 1 meant that a patient-important outcome was more likely to be specified as a primary outcome. We then combined ORs in a random-effects meta-analysis and calculated a pooled OR (along with its 95% CI and I^2 as a measure of heterogeneity) as an overall indicator of whether surgical trials uses patient-important outcomes as primary outcomes.

Exploratory meta-regression was modelled using the restricted maximum likelihood method in order to explore trial-level variables associated with the reporting of patient-important primary outcomes.

Results

Characteristics of included trials

We included 350 trials assessing a surgical intervention in our analysis (Fig. 1). Table 1 presents the characteristics of included trials. Most (335, 96%) were superiority trials, and parallel arm (331, 94·5%) trials, and 18% had an author with a reported epidemiology and/or statistics background. The primary outcome was specified in 225 (64%) trials.

A total of 8258 outcomes (4141 efficacy and 4117 harm outcomes) were reported in the included RCTs, with a median of 18 outcomes per trial. The mean proportion per

Table 1. Characteristics of included trials ($n = 350$)

Characteristic	No. (%)
Author with epi/stats degree	62 (18)
Study type	
Superiority/efficacy	335 (96)
Noninferiority/equivalence	15 (4)
Study design	
Parallel	331 (94.5)
Split body	17 (5)
Crossover	2 (0.5)
Journal impact factor	
None	23 (7)
< 1	41 (12)
1–2.4	131 (37)
2.5–4.9	112 (32)
5–9.9	30 (9)
≥ 10	13 (4)
Total sample size	
< 50	108 (31)
50–99	96 (27)
100–199	88 (25)
≥ 200	58 (17)
Multicentre	78 (22)
Trial registration	
Reported in text	57 (16)
Not reported but found online	37 (11)
Unclear	256 (73)
Primary outcome(s) specified	225 (64)
Generation of random sequence	
Adequate	147 (42)
Unclear	203 (58)
Allocation concealment	
Adequate	155 (44)
Inadequate or unclear	195 (56)
Blinding	
Any blinding	125 (36)
Patient	60 (17)
Carer	28 (8)
Outcome assessor	110 (31)
Primary outcome blinded	97 (28)
Unclear	225 (64)
Handling of attrition	
Intention to treat	93 (27)
Only follow-up reported	228 (65)
Inadequate	29 (8)
Source of funding	
Reported	153 (44)
Full industry	9 (6)
Partial industry	61 (40)
Nonindustry	49 (32)
No external source	34 (22)
Unreported	197 (56)

trial of patient-important outcomes reported was 60%, with slight variations when stratified for trial-level characteristics (Table 2). The most commonly reported patient-important outcomes were morbid events (29%), followed by intervention outcomes (10%), pain (7%) and function (7%). The mean proportions of surrogate outcomes and laboratory outcomes were 29% and 10%, respectively.

A median of 1 (interquartile range [IQR] 3) primary outcome was identified for each of the 225 trials that specified primary and/or secondary outcomes. Of these, 148 (66%) specified a patient-important outcome as a primary outcome. The most common patient-important primary outcome was a morbid event or symptom (92 trials, 41%), followed by intervention outcomes (25 trials, 11%) and function outcomes (24 trials, 11%). Seventy-four (33%) trials specified a surrogate outcome as a primary outcome, and 17 (8%) trials specified a laboratory outcome as a primary outcome (Table 3).

Are primary outcomes more likely to be patient-important?

Figure 2 depicts the results of our random-effects meta-analysis that pooled trial-level ORs of the association between patient importance and a primary outcome specification. Of the 225 trials that specified a primary outcome, 59 had entire rows or columns with zero cells in our 2 × 2 table and did not contribute to the meta-analysis. Patient-important outcomes were not associated with being a primary outcome (OR 0.82, 95% CI 0.63–1.07, I^2 = 21%). Exploratory meta-regression showed that trials that had an author with a declared epidemiology and/or statistics background had twice the odds of other trials of specifying a patient-important primary outcome (OR 2.08, 95% CI 1.10–3.94, p = 0.025, Table 4).

Table 2. Mean proportions of patient-important outcomes per trial, stratified by trial characteristics

	Trial characteristic; %*				
Outcome	All trials	Superiority trials	Noninferiority trials	Full or partial industry funding	Nonindustry funding
No. of trials	350	335	15	70	83
Any patient-important outcome	60	60	67	58	56
Mortality/survival	2	2	3	2	2
Pain	7	7	6	7	6
Function	7	7	10	10	6
Quality of life	4	3	8	5	6
Morbid events/symptoms	29	28	30	25	25
Patient satisfaction	2	2	0	0	3
Intervention outcomes	10	10	10	10	9
Surrogate outcomes	29	29	26	32	31
Laboratory outcomes	10	10	7	8	12

*Unless indicated otherwise.

Table 3. Proportion of trials reporting patient-important outcomes as primary outcomes, stratified by trial characteristics*

	Trial characteristic; no. (%)†				
Outcome	All trials	Superiority trials	Noninferiority trials	Full or partial industry funding	Nonindustry funding
No. of trials	225	212	13	58	58
Any patient-important outcome	148 (66)	139 (66)	9 (69)	35 (60)	34 (59)
Mortality/survival	10 (4)	10 (5)	0	0	2 (3)
Pain	20 (9)	20 (9)	0	4 (7)	4 (7)
Function	24 (11)	23 (11)	1 (8)	9 (16)	4 (7)
Quality of life	8 (4)	7 (3)	1 (8)	3 (5)	2 (3)
Morbid events or symptoms	92 (41)	86 (41)	7 (54)	20 (34)	22 (38)
Patient satisfaction	1 (0.4)	1 (0.5)	0	0	0
Intervention outcomes	25 (11)	25 (12)	0	7 (12)	8 (14)
Surrogate outcomes	74 (33)	71 (33)	3 (23)	25 (43)	20 (34)
Laboratory outcomes	17 (8)	15 (7)	2 (15)	5 (9)	3 (5)

*Data presented only from trials that explicitly specified a primary outcome. Some trials reported more than 1 primary outcome, hence the classifications were not mutually exclusive.
†Unless indicated otherwise.

Discussion

We systematically reviewed 350 recently published RCTs of surgical interventions to determine the extent of reporting of patient-important outcomes. We found that a mean proportion of 60% of outcomes per trial were patient-important, but substantial proportions of outcomes were surrogate (29%) or laboratory-based (10%). Of the surgical trials that specified a primary outcome, only two-thirds specified a patient-important primary outcome. Patient-important outcomes were not more likely to be specified as primary outcomes than surrogate or laboratory outcomes.

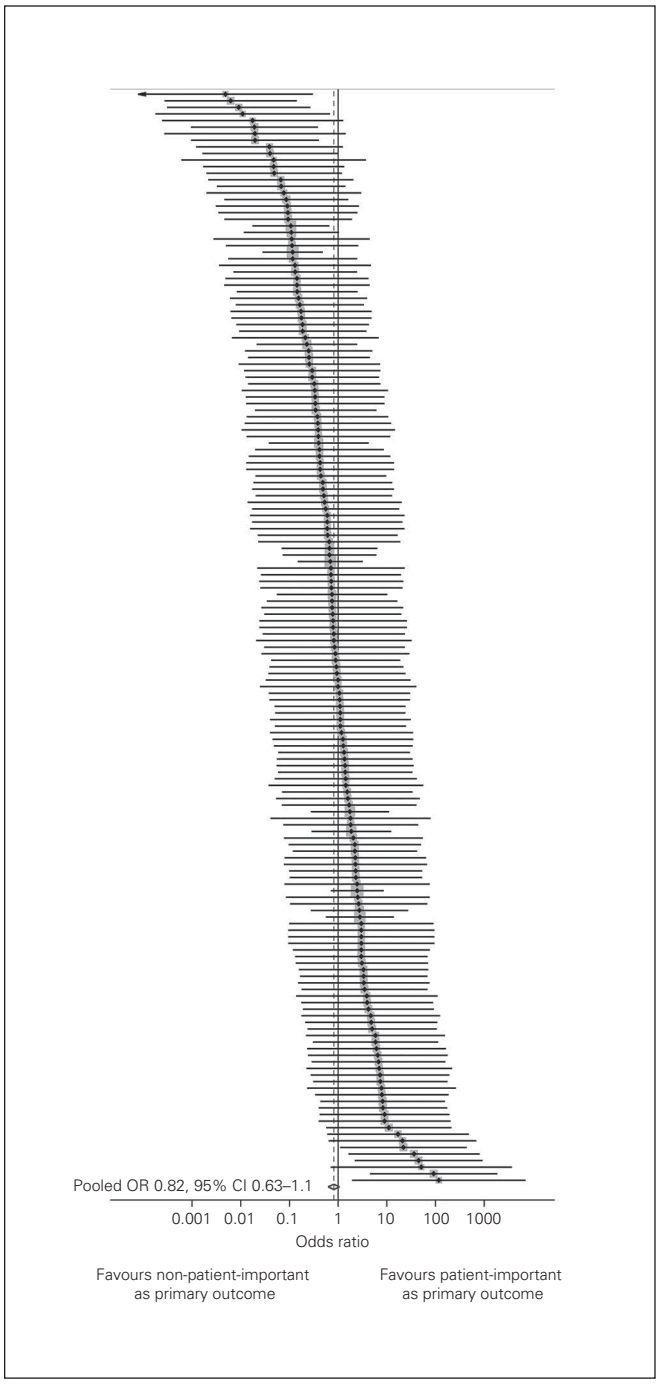

Fig. 2. Association between patient importance and specification as a primary outcome. Black points indicate odds ratios (ORs), grey squares indicate the relative weight of each trial, lines indicate 95% confidence intervals (CIs), and the diamond and dashed line the pooled OR.

Surrogate outcomes have been accepted as proxy measures of patient-important outcomes in clinical trials for many years. They are often easier and quicker to measure,[3] and statistical inferences can be made with smaller numbers of patients owing to larger treatment effect sizes.[17] Thus they are useful in the early evaluation of the bioactivity of an intervention,[2] and many drug interventions have been approved by regulatory bodies on the basis of a positive effect on surrogate outcomes.[18] However, assumptions are often made that surrogate outcomes lie on a causal pathway to a patient-important, clinically relevant outcome.[19] These assumptions of causality often rely on observational evidence, such as laboratory, ecologic and cohort studies,[3] and have resulted in grave misinterpretations of the benefits of some interventions. For example, class I antiarrhythmic medications were approved for use after myocardial infarction based on a proven reduction in ventricular ectopic beats (a surrogate outcome of mortality), but a subsequent large clinical trial[20] reported an increase in mortality with these agents, with thousands of patients likely to have been harmed (or killed) in the intervening period.[21] Hormone replacement therapy was shown to improve cholesterol levels in women, but subsequent evidence suggested an increase in the incidence of myocardial infarction and stroke with this therapy.[22] On the other hand, there are examples of trials for which surrogate and patient-important outcomes have been aligned, such as an increase in bone mineral density and fracture risk in trials of bisphosphonates.[23,24] Some commentators have argued strongly against the adoption of interventions based on surrogate outcomes.[2,18,19]

Surgical innovation has been criticized for not requiring rigorous evaluation before its availability to patients,[25] and it is likely that its early evaluation suffers from similar problems with the use of surrogate outcomes. We found that 60% of all outcomes were patient-important outcomes and that a slight majority (60%) of outcomes per study were patient-important. Further, 66% of trials specified patient-important primary outcomes. Studies in other specialty areas have found smaller proportions of patient-important outcomes in their samples of RCTs. Gandhi and colleagues[4] assessed RCTs of diabetes care and found that only 46% of trials reported any patient-important outcomes and only 18% of trials reported patient-important, primary outcomes. However, in their assessment of recent RCTs published in 6 high-impact general medical journals, Ciani and colleagues[17] found that 27% of trials specified a surrogate primary outcome. It is

possible that surgical interventions (compared with drug interventions) are more conducive to assessment by practical, clinically relevant outcomes, since surgeons often use patient performance and morbid events as markers of success. It is also likely that surgeons focus on clinically relevant outcomes (e.g., morbid/adverse events or functional outcomes) rather than laboratory measures to monitor their patients after surgery. Nevertheless, a significant proportion of outcomes (both primary and otherwise) in surgical trials remain non–patient-important, and this warrants some concern.

Strengths and limitations

Our study has a number of strengths. First, we used a protocol-driven systematic review design, which allows study replication and generalizability of the results to recently published surgical RCTs. Second, the definitions of patient-important outcomes were unambiguous, and there was minimal disagreement between the 2 researchers when the data were checked. This study also has weaknesses. First, it did not account for the occurrence of selective outcome reporting. It is likely that a certain proportion of outcomes remain unreported (possibly based on statistical significance). Further, some studies may have retrospectively selected primary outcomes based on their results or statistical significance, which would affect the analyses presented here on the association between primary outcomes and patient importance.[26] Second, to be included in the pooled analysis, a trial's outcomes had to populate a 2×2 table so that an OR could be calculated. Trials that had entire rows or columns with zero events (e.g., a trial reported entirely with surrogate outcomes), or trials that did not specify whether outcomes were primary or secondary, were excluded. This may have reduced the power of our analysis to detect a significant association, although the calculated CI was relatively narrow (0.63–1.07). Third, we used a trial sample published between August 2008 and May 2009. While this sample is now several years old, there was an inevitable time delay for collection and analysis of such a large amount of data, which was similar to previous studies.[27,28]

Table 4. Exploratory metaregression for the association between patient-important outcomes and specification as primary outcome

	Univariate		Multivariate*	
Variable	OR (95% CI)	*p* value	OR (95% CI)	*p* value
Industry funding	1.3 (0.6–2.7)	0.5	—	—
Author epi/stats degree	2.1 (1.1–4.0)	0.025	2.1 (1.1–4.0)	0.025
Noninferiority trial	0.5 (0.2–1.5)	0.2	0.6 (0.2–1.6)	0.3
Split-body/ crossover trial	2.1 (0.6–7.0)	0.2	2.0 (0.6–6.4)	0.3
Journal impact factor	1.0 (0.9–1.1)	0.8	—	—
Total sample size	0.99 (0.98–1.1)	0.2	0.99 (0.98–1.1)	0.6
Multicentre trial	0.9 (0.6–1.2)	0.5	—	—
Registered trial	1.0 (0.7–1.4)	0.9	—	—

CI = confidence interval; OR = odds ratio.
*Variables included in multivariate analysis only if $p < 0.25$ in univariate analysis.

Conclusion

We found that a substantial proportion of surgical trials did not specify a patient-important primary outcome and that patient-important outcomes were not more likely to be specified as primary outcomes than surrogate or laboratory outcomes. Consequently, many surgical trials may not be clinically useful. Surrogate outcomes should be used only when they have a strong, evidence-based link with a patient-important outcome. Trial reporting guidelines (e.g., CONSORT statement)[29] should include reporting on the clinical relevance of any surrogate outcomes measured, and trial funders and publishers should consider the importance of outcomes to patients in their decision-making.

Acknowledgements: The authors thank the librarians at the Cochrane Renal Group, Westmead, NSW, Australia, and the Ken Merten Library, Liverpool Hospital, Liverpool, NSW, Australia, for their assistance with the development of the search strategies used.

Affiliations: All authors are from the Whitlam Orthopaedic Research Centre, Ingham Institute for Applied Medical Research, Liverpool, New South Wales, Australia.

Funding: S. Adie is supported by scholarship grants from the National Health and Medical Research Council of Australia (Biomedical Postgraduate Scholarship), and the Royal Australasian College of Surgeons (Sir Roy McCaughey Research Fellowship). The funders had no role in the design, data collection or analysis of this study.

Competing interests: None declared.

Contributors: S. Adie and I. Harris designed the study. S. Adie, J. Naylor and R. Mittal acquired the data, which all authors analyzed. All authors wrote and reviewed the article and approved the final version for publication.

References

1. Korolija D, Wood-Dauphinee S, Pointner R. Patient-reported outcomes. How important are they? *Surg Endosc* 2007;21:503-7.
2. Fleming TR, DeMets DL. Surrogate end points in clinical trials: Are we being misled? *Ann Intern Med* 1996;125:605-13.
3. Bucher HC, Guyatt GH, Cook DJ, et al. User's guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring the effect of an intervention on surrogate end points. Evidence-Based Medicine Working Group. *JAMA* 1999;282:771-8.
4. Gandhi GY, Murad MH, Fujiyoshi A, et al. Patient-important outcomes in registered diabetes trials. *JAMA* 2008;299:2543-9.
5. Efficace F, Horneber M, Lejeune S, et al. Methodological quality of patient-reported outcome research was low in complementary and alternative medicine in oncology. *J Clin Epidemiol* 2006;59:1257-65.
6. Mathers SA, Chesson RA, Proctor JM, et al. The use of patient-centered outcome measures in radiology: a systematic review. *Acad Radiol* 2006;13:1394-404.

7. Baigent C, Blackwell L, Emberson J, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet* 2010;376:1670-81.
8. Ismail-Beigi F, Craven T, Banerji MA, et al. Effect of intensive treatment of hyperglycaemia on microvascular outcomes in type 2 diabetes: an analysis of the ACCORD randomised trial. *Lancet* 2010;376:419-30.
9. Giaccone G. Gefitinib in combination with gemcitabine and cisplatin in advanced non-small-cell lung cancer: a phase III trial — INTACT 1. *J Clin Oncol* 2004;22:777-84.
10. Barter PJ, Caulfield M, Eriksson M, et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med* 2007;357: 2109-22.
11. CAST Investigators. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *N Engl J Med* 1989;321:406-12.
12 Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Interm Med* 2009;151:264-9.
13. Lefebvre C, Eisinga A, McDonald S, et al. Enhancing access to reports of randomized trials published world-wide — the contribution of EMBASE records to the Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library. *Emerg Themes Epidemiol* 2008;5:13
14. Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002;31:150-3.
15. Chan A-W, 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.
16. Bradburn MJ, Deeks JJ, Berlin JA, et al. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med* 2007;26:53-77.
17. Ciani O, Buyse M, Garside R, et al. Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study. *BMJ* 2013;346:f457.
18. Moynihan R. Surrogates under scrutiny: fallible correlations, fatal consequences. *BMJ* 2011;343:d5160.
19. Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ* 2011;343:d7995.
20. Echt DS, Liebson PR, Mitchell LB. Mortality and morbidity in patients receiving encainide, flecainide, or placebo: the Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991;324:781-8.
21. Garratt C, Ward DE, Camm AJ. Lessons from the cardiac arrhythmia suppression trial. *BMJ* 1989;299:805-6.
22. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288:321-33.
23. Storm T, Thamsborg G, Steiniche T. Effect of intermittent cyclical etidronate therapy on bone mass and fracture rate in women with postmenopausal osteoporosis. *N Engl J Med* 1990;322: 1265-71.
24. Liberman UA, Weiss SR, Bröll J, et al. Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. *N Engl J Med* 1995;333:1437-43.
25. Barkun JS, Aronson JK, Feldman LS, et al. Evaluation and stages of surgical innovations. *Lancet* 2009;374:1089–96.
26. Hannink G, Gooszen HG, Rovers MM. Comparison of registered and published primary outcomes in randomized clinical trials of surgical interventions. *Ann Surg* 2013;257:818-23.
27. Chan A-W, Krleza-Jeric K, Schmid I, et al. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004;171:735-40.
28. Chan A-W, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005;330:753.
29. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869.