

primary outcomes reported in the ClinicalTrials.gov “primary outcome” field for 26 of 34 trials with an entry in that field.

Conclusions When authors fail to identify, register, and fully publish RCTs, then no protocol information beyond the abstract is available for systematic reviewers. Even if registered, information in ClinicalTrials.gov may disagree with that in the abstract. Protocols and amendments should be available at study inception to assist systematic reviewers.

¹Johns Hopkins Bloomberg School of Public Health, Mailroom W5010, 615 N Wolfe St, Baltimore, MD 21205, USA, e-mail: rscherer@jhsph.edu;
²Delta State University School of Nursing, Cleveland, MS, USA

SATURDAY, SEPTEMBER 12

Quality of Reporting I

Quality of Survey Reporting in High-Impact-Factor Journals

Carol Bennett, Sara Khangura, Jamie Brehaut, Ian Graham, David Moher, Beth Potter, and Jeremy Grimshaw

Objective Reports of survey research often lack information necessary for transparency and study reproducibility. Our objective was to identify a representative sample of published reports of survey research and evaluate them with respect to a broad range of reporting characteristics.

Design We purposively sampled the top 15 journals (by impact factor) from each of 4 broad health science topic areas (N = 60) where survey research is common: health science, public health, medicine, and medical informatics. We conducted an Ovid MEDLINE search using the terms “survey,” “questionnaire,” “health surveys,” and “data collection” to identify English-language studies published between January 2008 and February 2009. All citations were screened by 2 researchers to identify survey research that used a self-administered questionnaire as the primary data collection tool. Duplicate data abstraction employed a 32-item data collection tool designed to assess elements critical for transparency and reproducibility. These elements were identified through a comprehensive review of the literature identifying peer-reviewed survey reporting recommendations and from experts in survey research.

Results The search returned 1,719 citations resulting in 117 eligible studies. Preliminary results show that 13/117 (11%) described how representative the sample was of the population of interest and 47/117 (40%) discussed the generalizability of the results. With regard to reproducibility, 96/117 (82%) identified the mode of survey administration and 41/117 (35%) made the questionnaire used in the study available. Further analyses will outline the proportion of surveys that adequately report additional elements critical for transparency and reproducibility, such as a description of the survey’s development, data analysis, reporting of response rates and methods for calculation.

Conclusions Pilot data indicate that the quality of survey research reporting is suboptimal. The current work will help identify areas

where the development of an evidence-based reporting guideline would be expected to have the most impact on improving survey reporting.

Ottawa Health Research Institute, Clinical Epidemiology Program, ASB Box 693, 1053 Carling Ave, Ottawa, Ontario K1Y 4E9, Canada, e-mail: skhangura@ohri.ca

Citation of Prior Research in Reports of Clinical Trials

Karen A. Robinson¹ and Steven N. Goodman²

Objective Clinical trials should not be started or interpreted without consideration of prior trials that addressed the same or similar questions. Our objective was to systematically assess to what extent published reports of clinical trials cited relevant prior trials.

Design We searched Web of Science for 2004 combining terms for “meta-analysis” and “random” in title, abstract, and keywords. We used the meta-analyses to identify cohorts of RCTs addressing the same question. We then assessed, within each cohort, the extent to which trial reports cited the trials that preceded them. We calculated the proportion of prior trials that were cited (Prior Research Citation Index [PRCI]) and the proportion of the total available participant population cited (Sample Size Citation Index [SSCI]).

Results We identified 227 meta-analyses comprising 1523 trials in 19 disciplines. The median PRCI was 0.34 (lower decile 0.08, upper decile 0.76), meaning that only a third of relevant papers were cited. The median SSCI (0.44, lower decile 0.09, upper decile 0.8) was slightly larger than the PRCI, meaning that trials cited were a bit larger than trials not cited, but on average 56% of prior information was not referenced. Thirteen disciplines had PRCI of less than 40%. Of the 1101 RCTs that had 5 or more prior trials to cite, 511 (46%) cited either 0 or 1 prior trial.

Conclusions In reports of RCTs, for which the identification of prior research should be easier than with any other design, less than 40% of prior RCTs are cited, comprising less than 50% of the participants enrolled in all relevant prior trials. Further research is needed to explore the implications of this finding and the potential explanatory factors. The potential implications include ethically unjustifiable trials, wasted time and resources, incorrect conclusions, and unnecessary risks for trial participants.

¹Department of Medicine, Johns Hopkins University, 1830 East Monnet St, Room 8069, Baltimore, MD 21287, USA, e-mail: krobin@jhmi.edu;

²Departments of Oncology, Epidemiology, and Biostatistics, Johns Hopkins Schools of Medicine and Public Health, Baltimore, MD, USA.

Utility of Editorials and Commentaries That Accompany Publication of Research Studies

Diane Civic

Objective Editorials and commentaries that accompany the publication of research articles can enhance readers’ understanding of new studies. This analysis evaluates the extent to which these

editorials and commentaries provide information on potential biases and implications for practice beyond the material included in the study's discussion section.

Design Assessment of editorials and commentaries published in the same issue as research reports in *BMJ*, *JAMA*, *Annals of Internal Medicine*, *Lancet*, and the *New England Journal of Medicine* during 5 randomly selected months in 2008. Up to 3 editorials/commentaries per journal per month were included. Eligibility included linkage to a single randomized controlled trial (RCT), cohort study, or meta-analysis. A piloted structured data sheet was used to abstract information from commentaries and studies. For RCTs, risk of bias was evaluated using the Cochrane Collaboration tool.

Results Fifty-five editorial/commentary research study pairs were included. Description of the studies' main findings were concordant in 46 (84%) pairs. Authors of 32 studies (58%) and 23 commentaries (42%) reported competing interests. Twenty-two studies (40%) received industry funding; only 1 editorial/commentary discussed study sponsorship. Five editorials/commentaries (9%) mentioned strengths, and 12 commentaries (22%) mentioned limitations omitted by study authors. Thirty-three of the 41 RCTs in the sample (80%) met 1 or more Cochrane criteria for potential risk of bias. Thirteen of the 33 studies indicated these biases in their discussion sections, but only 5 editorials/commentaries mentioned them. Seventeen (31%) of the commentary/research study pairs had discordant recommendations regarding whether study findings warranted action (adopting or not adopting a practice).

Conclusions Editorials and commentaries did not routinely address strengths or weaknesses of empirical studies beyond those reported by study authors and rarely mentioned studies' links to industry. In about a third of the sample, editorials and commentaries provided a different perspective than study authors on whether findings were sufficient to recommend action.

Blue Cross/Blue Shield Association Technology Evaluation Center (BCBSA TEC), 225 North Michigan Ave, Chicago, IL 60601, USA, e-mail: diane.civic@bcbsa.com

Quality of Reporting II

Acknowledging Limitations in Biomedical Studies: The ALIBI Study

Milo A. Puhan,^{1,2} Nadine Heller,² Irena Joleska,² Lara Siebeling,³ Patrick Muggensturm,² Martin Umbehr,² Steve Goodman,¹ and Gerben ter Riet³

Objective To determine the proportion of clinical research papers acknowledging limitations, to categorize limitations, and to assess the degree of tempering of conclusions due to uncertainty arising from limitations.

Design Survey of medical research papers (November 2008 to February 2009). We included the first 10 papers describing randomized trials and observational or diagnostic studies published in 2007 in 30 journals (10 general medical and 20 specialty journals, of which half were first- and second-tier journals). Two reviewers independently evaluated the proportion and type of

acknowledged limitations and whether the wording of conclusions in the abstract and discussion section was tempered in light of limitations as perceived by the independent reviewers.

Results Seventy-three percent of the 300 papers acknowledged a median of 3 (range, 0-8) limitations in the discussion section, whereas 5.3% acknowledged a limitation in the abstract; 62.1% and 37.9% of acknowledged limitations referred to aspects of internal and external validity, respectively. Measurement errors (149) and selected study populations (115) were mentioned most frequently as limitations of internal and external validity, respectively. In 88.1% of the papers, tempering the conclusions because of limitations was not recognizable. Papers in general medical journals were more likely to acknowledge limitations at all (odds ratio, 2.27; 95% confidence interval [CI], 1.27-4.10) and in the abstract (OR, 3.57, 95% CI, 1.27-10.0), whereas the conclusions were not tempered more frequently (OR, 0.98, 95% CI, 0.43-2.33). First- and second-tier journals did not differ significantly (TABLE 12).

Table 12. Acknowledgment of Limitations in Medical Journals

Journal (10 Papers per Journal)	No. of Papers With Acknowledgment of Any Limitation	No. of Papers (With Acknowledgment of Limitations) With Tempering of Conclusions
<i>Annals of Internal Medicine</i>	10	3/10
<i>American Journal of Medicine</i>	10	1/10
<i>Annals of Family Medicine</i>	10	0/10
<i>Circulation</i>	10	1/10
<i>Heart</i>	10	2/10
<i>JAMA</i>	9	0/9
<i>Journal of Internal Medicine</i>	9	1/9
<i>Chest</i>	9	0/9
<i>Journal of Nuclear Medicine</i>	9	0/9
<i>European Journal Nuclear Medicine and Molecular Imaging</i>	9	1/9
<i>British Journal of Psychiatry</i>	9	1/9
<i>BMJ</i>	8	1/8
<i>Mayo Clinic Proceedings</i>	8	2/8
<i>Archives of General Psychiatry</i>	8	2/8
<i>Pediatrics</i>	8	2/8
<i>New England Journal of Medicine</i>	7	0/7
<i>Clinical Gastroenterology and Hepatology</i>	7	0/7
<i>Arthritis and Rheumatism</i>	7	1/7
<i>Diabetes Care</i>	7	1/7
<i>Diabetologia</i>	7	1/7
<i>Lancet</i>	6	2/6
<i>Medicine</i>	6	0/6
<i>Gastroenterology</i>	6	0/6
<i>American Journal of Respiratory and Critical Care Medicine</i>	6	1/6
<i>Journal of Rheumatology</i>	6	0/6
<i>Obstetrics & Gynecology</i>	6	2/6
<i>Pediatric Infectious Disease Journal</i>	5	0/5
<i>Surgery</i>	3	0/3
<i>Fertility and Sterility</i>	3	0/3
<i>Annals of Surgical Oncology</i>	1	1/1

Conclusions A limitation of our study is that acknowledged limitations could not be distinguished from true limitations. Limitations are acknowledged frequently in medical papers, but they are rarely reflected in abstracts or conclusions. As a consequence, readers may not fully realize the limitations of the findings reported. Our findings raise the suspicion that often limitations are acknowledged only pro forma and cannot play the crucial role in the scientific discourse they deserve.

¹Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Epidemiology Dept, Baltimore, MD 21205, USA, e-mail: mpuhan@jhsph.edu; ²Horten Centre for Patient-oriented Research, University of Zurich, Switzerland; ³Department of General Practice, University of Amsterdam, Amsterdam, the Netherlands

Reporting of Eligibility Criteria of Randomized Trials: Comparison Between Trial Protocols and Journal Articles

Anette Blümle,¹ Joerg J. Meerpohl,^{1,2} Gerd Antes,¹ and Erik von Elm³

Objective The reporting of randomized trials in journal articles strikes a balance between limited article length and completeness of study information. Comparing study protocols to subsequent publications, we aimed to study the frequency and nature of changes in eligibility criteria (EC) of randomized trials.

Design We established a cohort of protocols submitted in 2000 to the Research Ethics Committee of University Freiburg/Germany and subsequent full publications identified by electronic literature searches and survey of applicants. We identified 52 trial protocols with 78 publications. From protocols and publications we extracted information on EC differing between protocol and publication and classified them into 7 content categories. For each EC we examined whether it was added to the publication or missing, stated as inclusion or exclusion criterion, and whether the difference represented a minor or major change and would be suggestive of a smaller or larger study population.

Results The 78 publications were published in 50 journals, of which 23 (46%) endorsed the CONSORT statement in their author instructions (as of May 2009). For 1 trial, all the EC stated in the protocol matched with those reported in the publication. For 51 trials (98%) with 77 subsequent articles we found differences in EC reporting of different types. Of 1230 EC stated in protocols, 522 (42.4%) were matching between protocol and publication, and 708 (57.6%) were modified or missing in the publication (TABLE 13). A total of 572 EC (46.5%) were formulated as inclusion criterion, and 630 EC (51.2%) as exclusion criterion (28 labeled as “patient selection criteria”). The most frequent content categories of prespecified EC was comorbidity, medical treatment, and type/severity of illness. Seventy EC were new in the publications, for a total of 778 discordant EC. Most differences in EC between protocols and publications were deemed major; most of the published EC definitions were suggestive of larger study populations when the EC was prespecified but of smaller study populations when it was new.

Table 13. Characteristics of Eligibility Criteria in Trial Protocols and Publications

	EC Matching Between Protocol and Publication, No. (%)	EC Modified or Missing in Publication, No. (%)	Total EC Stated in Protocols, No. (%)	EC Added to Publication, No. (%)
Number	522 (100)	708 (100)	1230 (100)	70 (100)
Inclusion criterion	403 (77.2)	169 (23.9)	572 (46.5)	13 (18.6)
Exclusion criterion	98 (18.8) ^a	532 (75.1) ^b	630 (51.2) ^c	57 (81.4)
Content category				
Demographics	38 (7.3)	26 (3.7)	64 (5.2)	1 (1.4)
Type or severity of illness	104 (19.9)	100 (14.1)	204 (16.6)	9 (12.9)
Comorbidity	183 (35.1)	335 (47.3)	518 (42.1)	37 (52.9)
Treatment	92 (17.6)	156 (22.0)	248 (20.2)	16 (22.9)
Diagnostic procedures	26 (5.0)	8 (1.1)	34 (2.8)	...
Pregnancy-related issues	23 (4.4)	44 (6.2)	67 (5.4)	4 (5.7)
Other reasons	56 (10.7)	39 (5.5)	95 (7.7)	3 (4.3)
Minor/	...	125 (17.7)	...	13 (18.6)
Major difference between protocol and publication	...	583 (82.3)	...	57 (81.4)
Suggestive of a smaller/	...	49 (6.9)	...	60 (85.7)
larger	...	638 (90.1) ^d	...	8 (11.4) ^e
study population in publication compared to protocol				

^a21 of 522 (4.0%) labeled as “patient selection criteria”
^b7 of 708 (1.0%) labeled as “patient selection criteria”
^c28 of 1230 (2.3%) labeled as “patient selection criteria”
^d21 of 708 (3.0%) unclear
^e2 of 70 (2.9%) unclear

Conclusions Most articles do not mirror the exact definition of the trial’s study population as prespecified in the protocol. Because many users of trial information rely on data published in journal articles, the generalizability of trial results may be misinterpreted.

¹German Cochrane Centre, Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; ²Division of Pediatric Hematology & Oncology, Department of Pediatrics, University Medical Center Freiburg, Germany; ³Swiss Paraplegic Research, PO Box, CH-6207 Nottwil, Switzerland, e-mail: erik.vonelm@paranet.ch

Reporting of Continuous Outcome Measures in Randomized Clinical Trials: Is the Whole Story Being Told?

David L. Schriger,¹ Dan F. Savage,¹ and Douglas G. Altman²

Objective The CONSORT statement defines elements to be included in reports of randomized controlled trials (RCTs) but says little about the depiction of outcome data. In some reports of RCTs a minimal summary of the available data is presented; a 2-arm, 1000-patient trial might report only 2 means and 2 standard deviations. Such austere reduction may lead to misinterpretation of a trial. We designed our study to investigate the extent of such data reduction.

Methods We are evaluating 10 randomly selected RCTs with 1 or more continuous primary outcomes from 2007 to 2009 issues

of 20 leading medical journals. Using methods developed for the quantification of density of data in tables and graphs we measure the degree of data reduction in 2 ways. First, we note the format (text, table, or figure) that conveyed the most detailed information about the outcome and the way that information was conveyed (eg, mean alone; mean with standard deviation [SD], standard error of the mean [SEM], or confidence interval [CI]; histogram; or scatterplot). Second, we calculate the “percentage of available data presented” by dividing the number of data points and descriptive statistics presented for the outcome (the numerator) by the number of data points that could have been presented (the denominator) using a series of denominators of increasing stringency. We also calculate the percentage of data presented for the outcome that was best presented in each article.

Results In general, only a small fraction of available data are presented (mean for best outcome, 22%, median, 6%, range, 0.2%-100%). There was considerable heterogeneity by journal: mean range (2%-72%), median range (1%, 100%). For over half the journals the median percentage of data presented for the best outcome was under 10% and for 13 of 14 journals it was below 25%. The percentage of data presented for the best outcome was higher when presented in a figure (n = 49, mean 43%, median 22%), than a (n = 85, mean 10%, median 5%), or as text (n = 6, mean 7%, median 7%).

Discussion Reports of randomized trials present a small fraction of the available data. While the extent to which this leads to misinterpretation of trial results is unknown, scientific discourse would be enhanced by the presentation of the all of the data either in the paper or in online supplements.

¹University of California Los Angeles School of Medicine, Emergency Medicine Center, 924 Westwood Blvd, Suite 300, Los Angeles, CA 90024-2924, USA, e-mail: schriger@ucla.edu; ²Centre for Statistics in Medicine, University of Oxford, Oxford, UK

Quality of Reporting III

CONSORT for Improving the Quality of Reports of Randomized Trials: A Longitudinal Study of PubMed Indexed Articles

Sally Hopewell,¹ Susan Dutton,¹ Ly-mee Yu,¹ An-wen Chan,² and Douglas G. Altman¹

Objectives To examine the reporting characteristics and methodological details of randomized trials indexed in PubMed in 2000 and 2006 and to assess whether quality of reporting has improved following publication of the revised CONSORT Statement in 2001.

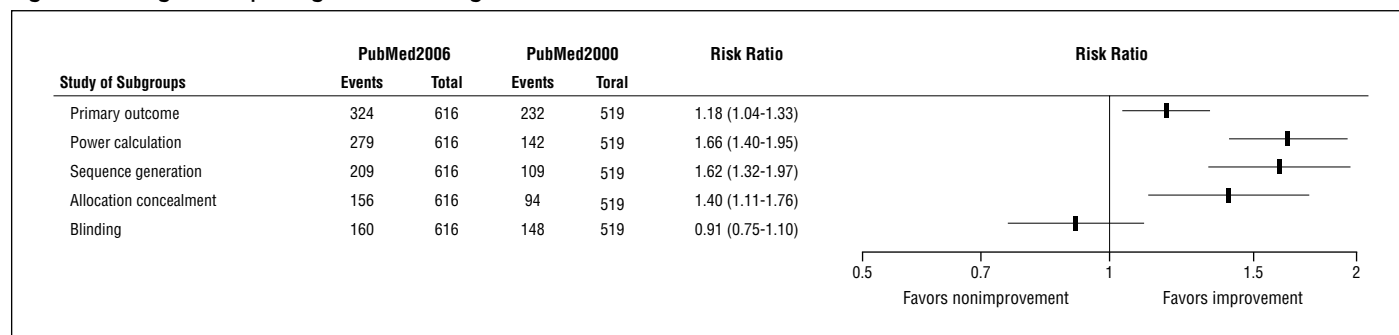
Design We examined all primary reports of randomized trials indexed in PubMed in December 2000 (n = 519) and December 2006 (n = 616). We included parallel-group, crossover, cluster, factorial, and split-body design studies; cost-effectiveness and diagnostic studies were excluded. We carried out single data extraction for a number of general and CONSORT specific items. Data were analyzed using STATA (ver 10); calculating the risk ratio (RR) (with 95% confidence intervals [CI]) to represent changes in reporting between 2000 and 2006.

Results The majority of randomized trials were 2-arm (73% in 2000 vs 76% in 2006), parallel-group trials (74% vs 78%), published in specialty journals (93% vs 90%), with a median sample size of 52 (interquartile range [IQR], 24-120) in 2000 and 62 (IQR, 33-152) in 2006. The proportion of drug trials decreased between 2000 and 2006 (76% vs 58%) and surgical trials increased (10% vs 21%). More articles reported details of the primary outcome (RR, 1.18; 95% CI, 1.04-1.33), power calculation (RR, 1.66; 95% CI, 1.40-1.95), random sequence generation (RR, 1.62; 95% CI, 1.32-1.97), and allocation concealment (RR, 1.40; 95% CI, 1.11-1.76) in 2006 (FIGURE 1). There was no significant difference in reporting of who was blinded (RR, 0.91; 95% CI, 0.75-1.10). In 2006, 28% of reports included a CONSORT flow diagram, and 61% gave the funding source; very few reported details of trial registration (9%) or access to the trial protocol (1%).

Conclusions Without important information about trial conduct it remains difficult to gauge the validity of trial results. Despite some progress in reporting of methodological details in recent years, there remains considerable room for improvement.

¹Centre for Statistics in Medicine, University of Oxford, Linton Road, Oxford, OX2 6UB UK, e-mail: sally.hopewell@csm.ox.ac.uk; ²Women’s College Research Institute, Department of Medicine, University of Toronto, Ontario, Canada

Figure 1. Changes in Reporting of Methodological Items Between 2000 and 2006



Blinding: Trial reports exactly who was blinded (eg, participants, care providers, outcome assessors).

CONSORT Guidelines for Reporting Abstracts of Randomized Trials: A Survey of Its Impact on High-Impact Journals

Sally Hopewell,^{1,2} Isabelle Boutron,¹ and Mike Clarke²

Objective To evaluate abstracts for reports of randomized trials published in 5 high-impact journals to assess the impact of CONSORT for Abstracts guidelines (published January 2008) and their influence on editorial policy.

Design We selected a random sample of 30 primary reports of randomized trials per journal per year from *Annals of Internal Medicine*, *British Medical Journal (BMJ)*, *Lancet*, *Journal of the American Medical Association (JAMA)*, and *New England Journal of Medicine (NEJM)* in 2007, 2008, and 2009, if indexed in PubMed with an electronic abstract. Secondary publications and economic analyses were excluded. Two authors extracted data independently using the CONSORT for Abstracts checklist. Data were analyzed using STATA (ver 10); 2007 and 2008 data are reported here, data for 2009 will be presented at the Congress.

Results A total of 284 abstracts were assessed (median participants per trial, 571 [interquartile range, 251 to 2005]). Most abstracts described the study as randomized in the title (216; 76%) and reported participant eligibility (256; 90%), interventions (218; 77%), objectives (274; 97%), primary outcome (201; 71%), result for each group with effect size (211; 74%), and precision (225; 79%). Allocation concealment (13; 5%), sequence generation (7; 2%), and specific details on who was blinded (12; 4%) were poorly reported as were trial design (65; 23%), funding source (3; 1%), harms (119; 42%), and number of participants randomized (137; 48%) and analyzed (92; 32%) in each group. There were substantial differences in the median proportion of CONSORT items reported across journals perhaps reflecting different editorial policies.

Conclusions Abstracts of randomized trials fail to meet a number of recommendations in the CONSORT for Abstracts guidelines. We hope the endorsement of the guidelines by the International Committee of Medical Journal Editors will herald improvements.

¹Centre for Statistics in Medicine, Wolfson College, Linton Road, Oxford, OX2 6UB UK, e-mail: sally.hopewell@csm.ox.ac.uk; ²UK Cochrane Centre, National Institute for Health Research, Oxford, UK

Reporting Clinical Trial Subgroup Analyses: A Proposal for Rigorously Assessing Heterogeneity in Treatments Effects

David M. Kent,¹ Peter M. Rothwell,² John P. A. Ioannidis,^{1,3} Douglas G. Altman,⁴ and Rodney A. Hayward⁵

Objective To develop a framework for the analysis and reporting of heterogeneity of treatment effect (HTE).

Design We reviewed the recent evidence on optimal statistical approaches to assessing HTE, and supplemented this with a systematic review of subgrouping practices in therapeutic trials that employ either clinical or FDA-accepted surrogate outcomes in 4 general medical journals (*BMJ*, *JAMA*, *Lancet*, and the *New England Journal of Medicine*).

Results Our initial review suggests that there is frequently tremendous variation in the baseline risk of the outcome of interest in clinical trial populations. These differences in risk may lead to clinically important HTE, such that the “average” benefit observed in the RCT summary result may be nonrepresentative of the treatment effect for many patients, including typical patients enrolled in the trial. Conventional subgroup analyses, which examine whether specific patient characteristics modify the effects of treatment, are usually unable to detect even large variations in treatment benefit (and harm) across risk groups because they do not account for the fact that patients have multiple characteristics simultaneously affecting outcome risk and potential for benefit. Risk-based subgroups using multivariate risk modeling are much better powered to detect HTE. Our systematic review of recently published clinical trials shows that this is often feasible but rarely done. While 61/93 (65.5%) studies reported subgroup analysis (a median of 4 per trial [range, 0-23; interquartile range, 2-6]), only 6 (6.5%) reported a risk-based subgroup analysis. Potentially applicable externally developed predictive models were available for 65 (69.9%) trials. Risk-based analysis was deemed feasible in all but 15 trials (16.1%).

Conclusions Trials that do not present interpretable absolute and relative treatment effects across risk categories are incompletely disclosing their results. Development of guidelines for subgroup analysis and the rigorous assessment of HTE using multivariable risk-based analysis could substantially improve the reporting of clinical trials.

¹Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, 800 Washington St, Box #63, Boston, MA 02111, USA, e-mail: dkent1@tuftsmedicalcenter.org; ²Department of Clinical Neurology, John Radcliffe Hospital, Oxford, UK, ³Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, ⁴Centre for Statistics in Medicine, University of Oxford, Oxford, UK, ⁵Division of General Medicine, Health Management and Policy, School of Public Health, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA

Reporting Guidelines for Clinical Research: A Systematic Review

David Moher,¹ Iveta Simera,² Kenneth Schulz,³ Donald Miller,⁴ Jeremy Grimshaw,¹ John Hoey,⁵ and Douglas G. Altman²

Objective To identify, document, and characterize all existing reporting guidelines developed to improve the quality of reporting of health research. A reporting guideline is a checklist, flow diagram, or explicit text to guide authors in reporting a specific type of research and is developed using explicit methodology, of which a consensus process is a crucial component.

Design A systematic review. We searched MEDLINE (February week 2 2009), EMBASE (2009 week 8), PsycInfo (February week 3 2009), and the Cochrane Methodology Register (2009 Issue 1). Two reviewers independently conducted a broad screen of all retrieved titles and abstracts and subsequently a full-text screen to determine final eligibility for those records passing the broad screen. To be included, the reporting guideline must have been developed using explicit methodology, including a consensus process, and be reported in English or French language. All

disagreements were resolved through consensus and, third-party arbitration, as needed. One researcher extracted descriptive information about each reporting guideline and the development process, using a recently developed checklist covering 4 phases of the development process. A pilot-tested, standardized data extraction form was used.

Results A total of 2,784 records were identified and 450 are currently being screened for final eligibility (87.0% agreement for broad screen). The majority of records were excluded because they were editorials or described clinical practice guidelines or reporting guidelines developed using a non-consensus-based process. A range of reporting guidelines have been identified related to clinical, laboratory, and economic health research. Some of the identified guidelines build on existing guidance (ie, for a specific clinical area) or are updates of previously published guidelines. A broad range of approaches were followed to develop reporting guidelines, but most included a face-to-face meeting of relevant stakeholders, including content experts, editors, and clinicians. Preliminary analysis indicates that reporting of specific elements of the guideline development process is suboptimal.

Conclusions This review helps to characterize similarities and differences across health research reporting guidelines. The diversity in development approaches and suboptimal reporting of the development process suggest there is a growing need to develop an instrument to help authors, editors, and others appraise the usefulness of any reporting guideline. The development of such a tool will be informed by this systematic review.

¹Clinical Epidemiology Methods Centre, Ottawa Health Research Institute, The Ottawa Hospital, General Campus, Critical Care Wing (Eye Institute), 6th Floor, 501 Smyth Rd, Ottawa, Ontario, Canada K1H 8L6, e-mail: dmoher@ohri.ca; ²Centre for Statistics in Medicine, Oxford University, Oxford, UK; ³Family Health International, Durham, NC, USA; ⁴University of Ottawa, Ottawa, Ontario, Canada; ⁵University of Toronto, Toronto, Ontario, Canada

Postpublication Citations, Indexing, Responses, and Online Publishing

Impact Factors of Secondary Journals

Cynthia Lokker, R. Brian Haynes, K. Ann McKibbin, and Nancy Wilczynski

Objective Secondary journals such as *Evidence-Based Medicine (EBM)*, *ACP Journal Club (ACPJC)*, and *Evidence-Based Nursing (EBN)* review more than 150 clinical journals and summarize articles that pass criteria for scientific merit and clinical relevance to practicing clinicians. Our objective was to calculate 2007 impact factors for the secondary journals in comparison with the article source journals.

Design A retrospective cohort study of articles abstracted in the secondary journals originally published in 2005 and 2006. We collected the number of citations in 2007 to these articles from the Institute for Scientific Information (ISI) Web of Science and calculated the 2007 impact factors for the secondary journals. We compared impact factors of the secondary journals with the pub-

lished impact factors of the journals represented within the secondary journals. We also compared the mean citations to summarized articles per journal to the published 2007 impact factors.

Results The 2005 and 2006 articles in the secondary journals were originally published in 84 journals, 82 with impact factors (median, 4.1; range, 0.85-52.9). The calculated impact factors for the secondary journals were 39.5 for *ACPJC*, 30.2 for *EBM*, and 9.3 for *EBN* (TABLE 14). Highest published impact factors for journals were *New England Journal of Medicine (NEJM)* (52.9), *Lancet* (28.6), *JAMA* (25.5), *Archives of General Psychiatry* (16.0), *Annals of Internal Medicine* (15.5), and *Journal of Clinical Oncology* (15.5). *ACPJC* and *EBM* had impact factors higher than all but *NEJM*. Of 100 journals categorized as “general and internal medicine” by ISI, the median impact factor was 1.3. Twelve journals had impact factors higher than *EBN* but none were nursing journals. Of the 46 nursing journals, the median impact factor was 0.9.

Table 14. Secondary Journal Impact Factors Based on Articles Originally Published in 2005 and 2006 and Abstracted in Evidence-Based Journals

Journal Name	Articles, No.	Citations in 2007, No.	Impact Factor	95% CI
ACPJC	286	11291	39.5	33.5-45.4
EBM	229	6918	30.2	23.5-36.9
EBN	189	1753	9.3	7.0-11.5

ACPJC, ACP Journal Club; EBM, Evidence-Based Medicine; EBN, Evidence-Based Nursing.

Conclusions The selection processes of evidence-based secondary journals identify articles at the time of publication that go on to garner more citations on average than other articles in the source publications. Whether this is simply due to selection processes or also due to stimulating citations of featured articles is unknown.

Health Information Research Unit, HSC 3H7, 1200 Main St W, McMaster University, Hamilton, Ontario L8N 3Z5, Canada, e-mail: lokker@mcmaster.ca

Does Journal Indexation Depend on the Origin of Publication? A Retrospective Cohort Study of Anesthesia Journals Indexed in MEDLINE and EMBASE

Martin Tramèr,¹ Nadia Elia,¹ Jean-Daniel Junod,¹ Teresa Dib,² and Christian Mazza³

Objective To study the association between the origin of publication of anesthesia journals and indexation rates in MEDLINE and EMBASE.

Design Retrospective cohort study of anesthesia journals published to 2005. Journals were systematically searched using International Standard Serial Number, US National Library of Medicine database, and Ulrich’s Periodicals Directory. We extracted information on origin of publication (United States [US], non-United States [non-US]), first and last date of publication, and indexation in MEDLINE and/or EMBASE. We computed indexation rates per 1000 journal years (IR) and rate ratios (RR) comparing IRs of US with non-US journals, both with 95% confidence intervals (CI).

Results We retrieved 325 journals, published from 1921 to 2005; 57 journals (25 US, 32 non-US) had been indexed in MEDLINE and 71 (21 US, 50 non-US) in EMBASE. In MEDLINE, IR of US journals was 20.8 (95% confidence interval [CI], 14.1-30.9), and the IR of non-US journals was 8.0 (95% CI, 5.6-11.3); RR 2.61 (95% CI, 1.48-4.55), $P < .001$. In EMBASE, IR of US journals was 16.9 (95% CI, 11.0-26.0), and the IR of non-US journals was 12.3 (95% CI, 9.3-16.2); RR 1.38 (95% CI, 0.79-2.34), $P = .223$. The RRs have significantly increased over time in MEDLINE (RR, 1.75 (95% CI, 1.07-2.88), $P = .024$; likelihood ratio test for interaction between origin and time period, $P = .003$, but not in EMBASE (RR, 1.27 (95% CI, 0.76-2.13), $P = .364$; likelihood ratio test for interaction, $P = .561$). Although IRs of US journals remained similar in both databases, IRs of non-US journals have dramatically decreased in MEDLINE only.

Conclusions MEDLINE and EMBASE both claim that journals are selected for indexation based on quality and are independent of origin. However, there is evidence that US anesthesia journals are significantly likelier to be indexed in MEDLINE than non-US anesthesia journals; this phenomenon is not found in EMBASE. Since both databases share very similar indexing criteria, quality of journals is unlikely to provide an alternative explanation to our findings.

¹Division of Anesthesiology, University Hospitals of Geneva, 24 rue Micheli-du-Crest, Geneva 14, 1211 Switzerland, e-mail: martin.tramer@hcuge.ch; ²University of Geneva, Switzerland; ³University of Fribourg, Fribourg, Switzerland

Authors' Reply to Substantive Criticism Raised in Online Letters to the Editor

Peter Götzsche,¹ Tony Delamothe,² Fiona Godlee,² and Andreas Lundh¹

Objective To study whether substantive criticism raised in online letters to the editor, defined as a problem that could potentially invalidate the research or make it less reliable than what it seemed to be, is addressed by the authors.

Design Cohort study of research papers published in the *BMJ* between October 2005 and September 2007. Two observers selected all papers where a substantive criticism was raised in the online Rapid Responses section, and 2 editors, blinded to the authors' replies, judged independently whether it was (1) minor, (2) moderate, or (3) major. Thereafter, the editors judged whether the criticism was (1) fully addressed, (2) partly addressed, or (3) not addressed by the authors. The criticism authors made the same judgment.

Results A substantive criticism was raised against 105 of 350 (30%) research papers, and the 2 editors judged it to be major in 34 and 54 cases, respectively. The authors had responded to 47 (45%) of the criticisms. The criticism was of similar severity in cases with and without authors' replies ($P = .72$). For the 47 criticisms with responses, we did not find a relation between the seriousness of the criticism and the adequateness of the replies, neither in the opinion of the editors ($P = .88$ and $P = .95$, $n = 47$) nor in the opinion of the criticism authors ($P = .83$, $n = 39$, 83% response rate). However, compared with the criticism authors, the editors

felt more often that the criticism was addressed (mean, 1.4 vs 2.3, $P < .001$, $n = 39$).

Conclusions Substantive criticism was common, but authors replied in only half of the cases. Editors judged the replies far more positively than the criticism authors. Resources permitting, editors might encourage authors to reply to substantive criticism and could aim for adequate replies, eg, by using the criticism authors as peer reviewers of the reply.

¹The Nordic Cochrane Centre, Rigshospitalet, Dept 3343, Blegdamsvej 9, DK-2100 Copenhagen, Denmark, e-mail: pcg@cochrane.dk; ²*BMJ*, London, UK

Reader Response to an Online Clinical-Decision Feature With Polls and Commenting

Patricia Kritek,^{1,2} Edward W. Campion,² and Pam Miller²

Objective Scientific, peer-reviewed journals are developing a greater presence online, but there are concerns about both the quality and the utility of the new, associated participatory features.

Design We assessed the use patterns for a free, interactive feature that includes a short case about a clinical controversy, expert commentary, a poll with 3 decision options, and an option to submit comments. All cases related to original research published at the same time. The topics ranged from outpatient management of asthma to end-of-life decision making.

Results An average of 6577 votes were cast (range, 3703-11205). The first case, mild-persistent asthma, had the most users (32822) with 18.5% voting. The voting rate increased over time to 37.5% for the last case. Participants were from 136 countries in the following regions: North America (63%), Europe (16%), Asia/Russia (9%), South America (9%), Australia/Oceania (2%), and Africa (1%). Those voting were physicians (85%), students (8%), other health professionals (5%), or other (2%). The 7 case-decision exercises received an average of 373 reader comments, although the number of comments correlated poorly with the number of votes cast, and only 6% of those who voted submitted a comment. In 7 cases the average number of comments published was 341. The greatest number of published submissions (492) was associated with a clinical scenario that received 6445 votes, while there were only 407 comments published for the feature with the most votes. Editors judged 92% of submitted comments appropriate for publication (range, 89%-95%).

Conclusions This new, participatory, online, clinical-decision feature at a medical journal's Web site evoked international reader responses, especially in the polls and, to a lesser extent, in the submission of comments. The vast majority of reader comments were judged appropriate for online publication although objective assessment of the quality of comments remains challenging.

¹Brigham & Women's Hospital, Pulmonary and Critical Care, 75 Francis St, Clinics Building, 3rd Floor, Boston, MA 02115, USA, e-mail: pkritek@partners.org; ²*New England Journal of Medicine*, Boston, MA, USA

**Use of the Internet by Print Medical Journals in 2003-2009:
A Longitudinal Observational Study**

David L. Schriger,¹ Rashida Merchant,¹ Ariana Chehrazhi,¹ and Douglas G. Altman²

Objective To examine the extent to which journals are augmenting print articles with online supplements and functions.

Design This is a longitudinal observational cohort study of 138 high-impact medical journals. The number and kinds of online supplements associated with each print article in a sample of 28 medical journals were assessed biennially starting with March 2003 issues using standardized abstraction forms. The use of “rapid response” pages that permit the public to provide post-publication review of papers was tracked for all journals that have this feature.

Results The number of journals providing online-only supplementary material increased from 32% (2003) to 50% (2005) to 61% (2007) to 64% (2009), and the percentage of articles that

contained supplementary material increased from 5% to 12% to 21% to 30%, respectively. This pattern was seen in both a random sample of journals (20) and a selective sample of journals (10) chosen because they were thought to have frequent online-only content. The number of video supplements also increased markedly from 2005 to 2007, showing a slight decrease in 2009. In contrast, the number of journals offering online postpublication review decreased from 12% (17/138) to 9% (12/138) from 2005 to 2007, and the percentage of articles with no responses was essentially unchanged at 82% (2005) and 81% (2007) (2009 data pending). **SEE TABLE 15.**

Conclusions The use of online-only articles and online-only supplements by print journals continues to increase. Postpublication critique of online-only articles provided by the journal does not seem to be taking hold.

¹University of California Los Angeles School of Medicine, Emergency Medicine Center, 924 Westwood Blvd, Suite 300, Los Angeles, CA 90024-2924, USA, e-mail: schriger@ucla.edu; ²Centre for Statistics in Medicine, University of Oxford, Oxford, UK

Table 15. Supplementary Web-Only Material to Print Journal Articles in 28 Journals 2003-2009

Journal ^a	No. of Articles ^b		Supplementary Text, %		Supplementary Tables and Figures, %		All Supplementary Material, %		Change, %	
	2003	2009	2003	2009	2003	2009	2003	2009	09 - 07	09 - 03
<i>Cancer Res</i> (r)	79	123	1	18	3	70	3	71	33	68
<i>Ann Intern Med</i> (r,h)	21	18	5	6	14	28	14	50	-3	36
<i>N Engl J Med</i> (h)	48	57	2	20	2	21	6	43	-1	37
<i>Circulation</i> (h)	85	85	1	25	1	29	1	41	16	40
<i>BMJ</i> (h)	71	66	20	29	22	30	26	30	4	4
<i>Intensive Care Med</i> (r)	26	23	0	22	0	22	0	26	15	26
<i>JAMA</i> (h)	38	32	0	3	0	6	0	25	-1	25
<i>Acad Emerg Med</i> (r)	17	12	0	25	6	8	6	25	6	19
<i>Osteoarthritis Cartilage</i> (r)	9	18	0	0	0	11	0	21	21	21
<i>Gastroenterology</i> (h)	29	43	0	8	0	18	0	20	6	20
<i>J Clin Psych</i> (r)	15	18	0	0	0	0	0	17	-2	17
<i>Rheumatology</i> (r)	15	22	0	5	0	9	0	14	4	14
<i>Trans Roy Soc Trop Med</i> (r)	26	16	0	0	0	13	0	13	7	13
<i>Mayo Clin Proc</i> (r)	15	14	0	7	0	7	0	11	11	11
<i>Anesth Analg</i> (r)	47	59	0	0	0	0	0	10	-8	10
<i>Am J Renal Physiology</i> (r)	20	24	0	0	0	8	5	8	-2	3
<i>Pain</i> (r)	23	24	0	0	0	8	0	8	3	8
<i>Lancet</i> (r,h)	87	58	0	17	1	5	2	5	-4	3
<i>Ann Emerg Med</i> (h)	18	17	0	0	0	0	0	0	-8	0
<i>J Am Acad Child Psychiatry</i> (r)	14	11	0	0	0	0	0	0	0	0
<i>Resuscitation</i> (r)	11	18	9	0	0	0	9	0	0	-9
Totals ^c (mean of row %s) No. of journals = 21	714	758	2	9	2	14	3	21	+5	+17
Random sample (r) No. of journals = 20	425	458	1	5	1	9	2	14	+4	+12
Handpicked sample (h) No. of journals = 10	397	376	3	11	4	14	5	21	+1	+17

^ar indicates journals selected at random; h, journals handpicked for sample; 5 journals in the random sample (*Arch Gen Psychiatry*, *Biol Blood Marrow Transplant*, *Ear Hear*, *J Assoc Res Otolaryngol*, *J Neurotrauma*) and 2 in the handpicked sample (*Ann Surg* and *Pediatrics*) had no supplementary material in any of the 4 years.

^bCounts do not include the 146 articles that were published solely online (87 of which were in *Pediatrics*).

^cThe number of articles reported in the final 3 rows is based on the relevant journals from the 21 listed in the table. The entries for the other columns in these rows are calculated as the average of the percentages for each journal in that sample.