

Reviewing the reviewers: the quality of reporting in three secondary journals

P.J. Devereaux, Braden J. Manns, William A. Ghali, Hude Quan, Gordon H. Guyatt

Abstract

Background: Secondary journals such as *ACP Journal Club* (ACP), *Journal Watch* (JW) and *Internal Medicine Alert* (IMA) have enormous potential to help clinicians remain up to date with medical knowledge. However, for clinicians to evaluate the validity and applicability of new findings, they need information on the study design, methodology and results.

Methods: Beginning with the first issue in March 1997, we selected 50 consecutive summaries of studies addressing therapy or prevention and internal medicine content from each of the ACP, JW and IMA. We evaluated the summaries for completeness of reporting key aspects of study design, methodology and results.

Results: All of the summaries in ACP reported study design, as compared with 72% of the summaries in JW and IMA ($p < 0.001$). In summaries of randomized controlled trials the 3 secondary journals were similar in reporting concealment of patient allocation (none reported this), blinding status of participants (ACP 62%, JW 70% and IMA 70% [$p = 0.7$]), blinding status of health care providers (ACP 12%, JW 4% and IMA 4% [$p = 0.4$]) and blinding status of judicial assessors of outcomes (ACP 4%, JW 4% and IMA 0% [$p = 0.4$]). ACP was the only one to report whether investigators conducted an intention-to-treat analysis (in 38% of summaries [$p < 0.001$]), and it was more likely than the other 2 journals to report the precision of the treatment effect (as a p value or 95% confidence interval) (ACP 100%, JW 0% and IMA 55% [$p < 0.001$]).

Interpretation: Although ACP provided more information on study design, methodology and results, all 3 secondary journals often omitted important information. More complete reporting is necessary for secondary journals to fulfill their potential to help clinicians evaluate the medical literature.

Busy clinicians never have the time to screen all of the journals or to read all of the articles relevant to their practice.¹⁻³ Secondary journals offer a potential solution for clinicians striving to keep abreast of important studies. They screen a number of medical journals and summarize information from the most important and relevant articles in condensed 1- to 2-page summaries. The popularity of secondary journals is growing, and some physicians are starting to rely on them as a primary source of medical information.⁴

To apply the results of clinical studies to their practice, experienced clinicians often need only a few key details of study methodology and results to assess the likelihood that the results will be unbiased (the study validity) and to assess the magnitude and precision of the treatment effect.^{5,6} Because the nature and quality of reporting in secondary journals remains largely unexplored, we evaluated the extent to which 3 popular secondary journals include key aspects of study design, methodology and results in their summaries.

Methods

For the purposes of this observational study we defined a secondary journal as one that publishes abstracts or summaries, or both, of studies previously reported in other journals. We will refer to these abstracts and summaries, which can consist of a structured abstract,

Research

Recherche

At the time of writing, Dr. Devereaux was with the Department of Medicine, Dalhousie University, Halifax, NS; he is now with the Department of Medicine, McMaster University, Hamilton, Ont. Drs. Manns and Ghali are with the Departments of Medicine and Community Health Sciences, University of Calgary, Calgary, Alta. Dr. Quan is with the Department of Community Health Sciences, University of Calgary, Calgary, Alta. Dr. Guyatt is with the Departments of Medicine and of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ont.

This article has been peer reviewed.

CMAJ 2001;164(11):1573-6

See related article page 1580

[Return to May 29, 2001 Table of Contents](#)

commentary or critique, as “summaries.” The 3 secondary journals selected for our study were the *ACP Journal Club* (ACP), *Journal Watch* (JW) and *Internal Medicine Alert* (IMA). We chose these journals because they are widely read by internists and have a relatively large circulation (ACP 85 000, JW 30 000 and IMA 5000).

We included summaries of studies addressing questions of therapy or prevention and internal medicine content (i.e., any research in adult or adolescent medicine). We excluded summaries of meta-analyses and review articles. We included the first 50 consecutive summaries in each of the 3 secondary journals that fulfilled the eligibility criteria, starting with the first publication in March 1997. Two of us (W.A.G. and G.H.G.) independently evaluated all of the summaries in the 3 journals for eligibility and resolved disagreements by consensus in subsequent discussions. The chance-correct agreement for this process, assessed through means of a kappa statistic, was 0.82 (95% confidence interval [CI] 0.75–0.88).

We evaluated all information included in the selected summaries. We determined whether each summary provided an explicit statement that allowed the reader to classify the study as a randomized controlled trial (RCT) or an observational study (e.g., cohort or case-control study). We also determined whether summaries of RCTs provided the following information: concealment of patient allocation (through explicit mention of concealment, lack of concealment, or a description of a method of concealment or non-concealment as described by Schulz and associates⁷); method of analysis (intention to treat or other); blinding status of participants, health care providers and judicial assessors of outcomes (individuals who ultimately decide whether a patient meets criteria for the outcome being evaluated); proportion of participants lost to follow-up; *p* value or 95% CI for the main outcome; and, for summaries of RCTs with dichotomous outcomes, the proportion of participants in each treatment arm with the outcome of interest and the associated relative risk, relative risk reduction, absolute risk reduction, number needed to treat and number needed to harm. Two of us (P.J.D. and B.J.M.) independently assessed the presence or absence of each feature in the summaries; the first 17 summaries were reviewed by both abstractors, after which a high level of agreement was found, and the remaining summaries were evenly divided between the 2 abstractors for review. Because the 3 secondary journals use characteristic formats, we could not blind the abstractors to the secondary journal. The interobserver kappa statistic for the 17 summaries assessed by both abstractors was 0.64 (95% CI 0–1.00) for the reporting of the number needed to treat and the number needed to harm, and ≥ 0.80 (95% CI 0.54–1.00) for the reporting of all other variables.

The main analysis compared the frequency with which the summaries reported key aspects of study design, methodology and results. We chose a χ^2 or Fisher's exact test, depending on cell size, to test the statistical significance of differences across secondary journals.

Results

We screened 71 summaries in ACP, 117 in JW and 131 in IMA to identify the 50 summaries per journal that fulfilled our eligibility criteria. The full-text articles summarized by ACP, JW and IMA were originally published in 14, 22 and 22 medical journals, respectively.

Study design was reported in 100% of the ACP summaries, as compared with 72% of the summaries in JW and IMA ($p < 0.001$). This analysis also revealed that ACP summarized only RCTs. Of the JW and IMA summaries 46% were of RCTs, 26% were of observational studies, and 28% were of studies in which the design was not clarified.

Review of the secondary journals' reporting of RCT methodology revealed infrequent presentation of key methodological factors (Table 1). None of the summaries mentioned concealment of patient allocation, and few mentioned the blinding status of health care providers and of judicial assessors of outcomes. Although all 3 secondary journals reported more frequently the blinding status of participants, over 30% of the summaries omitted this detail. The ACP summaries reported more frequently than the other journals' summaries whether the analysis was intention to treat and what the proportion was of participants lost to follow-up, but the frequency was still low (i.e., less than 50% of the time).

Table 2 presents the reporting of results in the summaries of RCTs that had dichotomous outcomes (28 summaries in ACP, 15 in JW and 11 in IMA). All 3 journals frequently reported the proportion of participants in each study arm with the outcome of interest. ACP reported the relative risk, relative risk reduction, absolute risk reduction, number needed to treat and number needed to harm far more frequently than did JW or IMA. Across all studies, ACP reported on a measure of the precision of the result

Table 1: Frequency with which 3 secondary journals reported methodological factors in summaries of randomized controlled trials

Factor	Journal; % of summaries			<i>p</i> value
	ACP	JW	IMA	
Concealment of patient allocation	0	0	0	1.0
Intention-to-treat analysis	38	0	0	< 0.001
Blinding status				
Participants	62	70	70	0.7
Health care providers	12	4	4	0.4
Judicial assessors of outcomes	4	4	0	1.0
Proportion of participants lost to follow-up	46	13	0	< 0.001

Note: ACP = *ACP Journal Club*, JW = *Journal Watch*, IMA = *Internal Medicine Alert*.

(*p* value or 95% CI) in all of its summaries, whereas IMA did so 55% of the time and JW never did so (*p* < 0.001).

Interpretation

There are 2 ways in which secondary journals can make it easier for clinicians to use the medical literature to solve patient problems. First, they can report only relevant and methodologically strong articles. Second, they can use an abbreviated format to present key information. We assessed the performance of 3 secondary journals in the second of these roles. We specifically evaluated the reporting of information that is crucial for determining the validity of a study and for understanding the results sufficiently well to apply them to patient care.

With regard to RCTs, a number of studies have shown that inadequate concealment of patient allocation, such that those responsible for enrolling patients are aware of the arm to which the patients will be allocated if enrolled, may lead to overestimation of treatment effects,⁷⁻¹⁰ as may lack of blinding.⁷ Thus, concealment of patient allocation and blinding of participants, health care providers and those assessing outcomes are critical methodological factors that are necessary for assessing a study's validity. An intention-to-treat analysis includes participants in the group to which they were allocated, irrespective of whether they received the prescribed treatment. Failure to analyze by intention to treat defeats the purpose of randomization and may bias the results.¹¹⁻¹⁴ Reporting the proportion of participants lost to follow-up is a final factor that we believe bears strongly on the likelihood studies will produce an unbiased estimate of the treatment effect.^{11,14}

Of these determinants of validity, ACP reported the study design in all of the summaries, the blinding status of participants in 62%, the proportion of participants lost to follow-up in 46%, and other criteria in less than 40%. Where the ACP summaries differed from those in JW and IMA, the reporting in the latter 2 journals was less complete. The most significant omissions related to study design, which JW and IMA failed to specify in 28% of their summaries. Thus, all 3 secondary journals failed to provide much of the information readers need to assess study validity.

Because treatment decisions inevitably involve trade-offs

between risks and benefits, clinicians require information not only about validity and whether treatment is effective, but also about the magnitude and precision of the estimate of the treatment effect. ACP provided some estimate of the magnitude of effect (relative risk reduction, risk difference or number needed to treat) in most of its summaries, whereas JW and IMA seldom provided these data. ACP always provided either a *p* value or confidence interval, information that JW omitted in its summaries and IMA provided in just over 50% of its summaries.

To our knowledge our study is the first to evaluate the quality of reporting in secondary journals. Previous work has demonstrated suboptimal reporting of RCTs in full-text journals.¹⁵⁻¹⁷ Those results, along with our findings, indicate the need for enhanced reporting of primary full-text articles and of secondary summaries of these studies.

For our analysis we assumed that the goal of secondary journals is to provide information that clinicians can apply directly to their clinical practice, thus enhancing their efficiency in using the original medical literature to guide patient care decisions. Secondary journals may have different goals. Indeed, JW includes a statement that its summaries are not intended for use as the sole basis for clinical treatment nor as a substitute for reading the original journal articles. These statements suggest that JW has different goals, such as simply alerting clinicians to information that they may want to explore further.

Although clinicians have not been surveyed on how they use secondary journals, we find it implausible that many use secondary journals largely as a stimulus to seek out and read original journal articles. For those who do use secondary journals in this way, having more informative summaries that identify the key methodological factors and results will aid in selecting which studies to spend time retrieving and reviewing. Finally, and most important, we would argue that, by not aiming to provide complete enough information to guide clinical practice directly, secondary journals are abandoning their most important potential role. At the same time, we acknowledge that clinicians' views of the optimal goals of secondary journals will differ with their values and are a matter for debate.

Whatever the goals of secondary journals, they should be able to provide brief and complete summaries. For example,

Table 2: Frequency with which the secondary journals reported results in summaries of randomized controlled trials with dichotomous outcomes

Type of result	Journal; % of summaries			<i>p</i> value
	ACP	JW	IMA	
Proportion of participants in each study arm with the outcome of interest	100	87	82	0.06
Relative risk, relative risk reduction or absolute risk reduction	86	0	27	< 0.001
Number needed to treat or number needed to harm	71	0	9	< 0.001

consider the statement, "This concealed placebo-controlled RCT, with effective blinding of participants, health care providers and judicial assessors, 99% of participants available for follow-up and an intention-to-treat analysis, assessed the effect of amiodarone in patients with heart failure." A second sentence could include all of the crucial data about the magnitude and precision of treatment effects.

Our study has limitations. We focused only on trials addressing therapeutics and prevention and therefore were unable to comment on the reporting of summaries that reviewed studies of diagnosis, prognosis or harm. Our evaluation of reporting methods and results focused on RCTs. Further work is needed to evaluate summaries of observational studies. The abstractors were aware of the secondary journal that published the summary they were assessing. However, the minimal interobserver variation in abstracting information suggests that it was unlikely that this lack of blinding influenced our findings. Finally, we evaluated only 3 secondary journals; however, we believe that most other secondary journals are no more likely than these 3 to report study design, methods and results. As such, our results may be widely generalizable.

One of us (G.H.G.) is an associate editor of ACP. Bearing in mind this possible conflict of interest, we were scrupulous in selecting criteria for optimal reporting and in conducting our assessment. Furthermore, we have presented a detailed rationale for our choice of factors that are crucial to allow clinicians to evaluate the validity and applicability of study results. Readers must decide whether our methods withstand the more intense scrutiny that is appropriate whenever issues of conflict of interest arise.

We acknowledge that brief summaries in secondary journals can never be substitutes for full-text reports, that detailed review of methodology will always raise additional issues and that, on occasion, those issues will have important implications for study validity and applicability. For instance, the term "intention-to-treat analysis" suffers from ambiguity and variability in interpretation. Nevertheless, few clinicians have either the time or the skills to conduct the detailed review required to elucidate such issues. Secondary journals, if they scrupulously report methodological details using the most transparent terms available, can provide summaries that, although not perfect, can serve clinicians well.

For those who believe that secondary journals should provide summaries that clinicians can apply directly to patient care, our results have a clear message. By implementing a systematic and easily achieved approach to reporting a small number of key features of methods and results, secondary journals could fulfil their potential to help clinicians deliver efficient, evidence-based care. For those who see other goals for secondary journals, our results are also im-

portant. First, they suggest the need for an explicit formulation of secondary journals' goals, and a debate about what the optimal goals might be. Second, they suggest the need for primary research on how, in an era of increasing pressures and time constraints, clinicians actually use information from secondary journals.

Competing interests: None declared for Drs. Devereaux, Manns, Quan and Ghali. Dr. Guyatt is associate editor of *ACP Journal Club*.

Contributors: Dr. Devereaux was the principal investigator of the study; he proposed the study, collected data, was involved with data interpretation and led the writing of the manuscript. Drs. Manns, Ghali, Quan and Guyatt contributed to the study design, the collection and interpretation of the data, and the revising of the manuscript. Drs. Ghali and Quan undertook the data analyses.

Acknowledgements: Dr. Devereaux is supported by a Heart and Stroke Foundation of Canada/Canadian Institutes of Health Research Fellowship Award. Dr. Manns is supported by a Kidney Foundation of Canada/Alberta Heritage Foundation for Medical Research Fellowship Award. Dr. Ghali is supported by a Population Health Investigator Award from the Alberta Heritage Foundation for Medical Research and holds a Government of Canada Research Chair in health services research.

References

- Davidoff F, Haynes B, Sackett D, Smith R. Evidence based medicine. *BMJ* 1995;310:1085-6.
- Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based medicine: how to practice and teach EBM*. New York: Churchill Livingstone; 1997. p. 8-9.
- Lundberg GD. Perspective from the editor of *JAMA*, the *Journal of the American Medical Association*. *Bull Med Libr Assoc* 1992;80:110-4.
- Naylor CD. Where's the meat in clinical journals [letter]? *ACP J Club* 1994;120 (May-June):87.
- Ad Hoc Working Group for Critical Appraisal of the Medical Literature. A proposal for more informative abstracts of clinical articles. *Ann Intern Med* 1987;106:598-604.
- Oxman AD, Sackett DL, Guyatt GH, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. I. How to get started. *JAMA* 1993;270:2093-5.
- Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
- Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 1998;317:1185-90.
- Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983;309:1358-61.
- Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;352:609-13.
- Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? *JAMA* 1993;270:2598-601.
- Greenhalgh T. How to read a paper, assessing the methodological quality of published papers. *BMJ* 1997;315:305-8.
- Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int J Epidemiol* 1992;21:837-41.
- Jadad A. *Randomised controlled trials*. London: BMJ Books; 1998. p. 35-6.
- Sonis J, Joines J. The quality of clinical trials published in the *Journal of Family Practice*, 1974-1991. *J Fam Pract* 1994;39:225-35.
- Ah-See KW, Molony NC. A qualitative assessment of randomized controlled trials in otolaryngology. *J Laryngol Otol* 1998;112:460-3.
- Bath FJ, Owens VE, Bath PMW. Quality of full and final publications reporting acute stroke trials: a systematic review. *Stroke* 1998;29:2203-10.

Reprint requests: Dr. P.J. Devereaux, Department of Clinical Epidemiology and Biostatistics, Rm. 2C12, Faculty of Health Sciences, McMaster University, 1200 Main St. W, Hamilton ON L8N 3Z5; philipi@mcmaster.ca