

## Poster Session Abstracts

### Advertising

#### A Survey on Publication Standards of Medical Drug and Device Advertisements Published in Core Medical Journals in China

Liang DU,<sup>1</sup> Yao-long CHEN,<sup>1</sup> Min CHEN,<sup>2</sup> and You-ping LI<sup>1</sup>

**Objective** Medical drugs and devices are the most widely used products to provide medical service in any country. Strict pre- and postmarketing assessments are needed during their whole lifespan. More than 6,000 manufacturers exist in China due to the huge market and large profits. Most of them produce the same products but publish numerous different advertisements for their limited products. There are more than 1,000 medical journals in China, and most of them will publish advertisements. Presumably, advertisements in core medical journals have far more influence than ones in other media due to the journal's academic reputation, but they often lack standards or evidence to judge the quality of the advertised products. The Standards for the Examination and Publication of Drug and Device Advertisements were implemented in 1995 and updated in 2007 and 2009. The 2 standards stated that only advertisements with both advertisement license and production license numbers could be published. This study aims to learn the current status, publication standards, formats, and contents of medical advertisements published in journals in China and to discuss the possibility of evidence-based evaluation and standards.

**Design** We reviewed issue 1 of 222 core medical journals published in 2008 and indexed by *A Guide to the Core Journals of China* (2004 version), the most important database to index the top 20% academic journals, to identify basic journal information and the content of their advertisements. The general and trade name of the drugs and the advertisements and production license number of the drugs and devices were collected. We used EXCEL software for data input and SPSS 13.0 for statistical analyses.

**Results** Two hundred eighteen journals were hand-searched and evaluated. The other 4 journals were excluded because a print version could not be found. A total of 1201 advertisements were published in 159 (72.9%) of the journals, with an average of 5.5 advertisements (range, 1-37) per journal. Of the advertisements, 910 (75.8%) were related to medical drugs or devices, including 598 (49.8%) drug and 312 (26%) medical device advertisements. Most advertisements were published in clinical and specialty medical journals. A total of 518 (86.6%) drugs advertisements had both advertisement license and production license number but only 116 (36.1%) of medical devices advertisements stated the advertisement license and production license number. References were found in less than 10% of advertisements.

**Conclusions** The medical drug advertisements published in core medical journals of China lack sufficient publication standards, and medical devices advertisements are even worse. We cannot assess the efficacy, safety, and cost-effectiveness of advertisement production according to the currently limited, unclear, and highly commercialized advertisements. It is necessary to improve the publication standard for advertisements so that they provide enough

necessary evidence, develop proper format and approach, and enhance their application and management. We are conducting a further survey on the information of advertising products related to clinical trial registration, publication, and evidence grade to clearly indicate their real efficacy, safety, and cost-effectiveness.

<sup>1</sup>Chinese Cochrane Centre, West China Hospital of Sichuan University, No. 37 Guoxuexiang, Chengdu 610041, China, e-mail: yzmylab@hotmail.com; <sup>2</sup>Department of Traditional Chinese Medicine, No. 3 People's Hospital, Chengdu, China

#### Papers Supporting Advertisements in Medical Journals: A Continuation of the Case-control Study

Vasily Vlassov

**Objective** Print journals depend on advertisements. To describe the patterns of the articles published in support of advertisements, including types of articles, differences in the practice of advertisers, and time trends.

**Design** Case-control study, extending work from a previous study. From a convenience sample of 2 international and 3 Russian peer-reviewed journals, 3 were selected for the extension of the study from 2000 to 2008 because of availability. Only advertisements for medical products, not education and jobs, were considered. The connection of the article to the advertisement was judged by the presence in the article of a positive comment on the advertised product or the use of the trade name and/or manufacturer.

**Results** Three journals have a rather constant level of association of published advertisements with the content of a journal (odds ratio, 2.2-40 during different years) and interleaf advertisements with articles. All major advertisers have a similar level of association of their advertisements with the content of the journals. Some advertised products do not show the statistically significant association with articles published in the same issue.

**Conclusions** Manipulation of the journal content for support of the advertisements is a long-lasting practice. The stable level of association during 8 study years means that this practice is not a "period effect." Association is visible also from juxtaposition of advertisements and related articles. The strength of the association and use of this instrument by all major advertisers means that part of the "scientific" content of the journals is misleading. In journals successfully attracting advertisements, this misleading content may be included in up to 50% of published articles.

Moscow Medical Academy, POB 13, Moscow, 109451, Russia, e-mail: vlassov@cochrane.ru

## Authorship and Contributorship

### Factors Associated With Multiple Authorship in Peer-Reviewed Papers Regarding Pregnancy Over 3 Decades

Brian Mercer,<sup>1</sup> Catherine Spong,<sup>2</sup> and James Scott<sup>3</sup>

**Objective** To determine if variation in authorship number in peer-reviewed papers regarding pregnancy can be attributed to differences in journal type, funding, or multicenter studies.

**Design** Using PubMed we identified original studies with abstracts and the key-word “pregnancy” published from 1975 to 2008 in 3 obstetric specialty journals: *Obstetrics & Gynecology (OG)*, *American Journal of Obstetrics & Gynecology (AJOG)*, and *British Journal of Obstetrics & Gynecology (BJOG)* and 3 general medical journals: *New England Journal of Medicine (NEJM)*, *Journal of the American Medical Association (JAMA)*, and *Lancet (LT)*. Univariable and multivariable comparisons were performed for author number, more than 6 (GT6) and more than 10 (GT10) authors, corporate authorship (none specified), and group authorship (authors plus group), according to publication year, journal type, funding, and multicenter study.

**Results** Of 12,981 papers, GT6 authorship occurred in 15.7% (*OG*, 7.1%; *AJOG*, 17.0%; *BJOG*, 11.8%; *JAMA*, 33.9%; *NEJM*, 39.3%; *LT*, 31.7%;  $P < .0001$ ). GT6-authorship was more common in general medical journals (34.3% vs 13.3%; odds ratio [OR], 3.40, 95% confidence interval [CI], 3.02-3.83), multicenter (45.3% vs 13.5%; OR, 5.29 [4.59-6.09]), and funded (18.1% vs 9.0%; OR, 2.24 [1.97-2.55]) studies. From 1975 to 2008, GT6-authorship increased from 0% to 72.7% in general medical and 1.7% to 30.7% in obstetric journals, each  $P < .0001$ . The difference in GT6 authorship between journal types increased over time,  $P < .001$ . GT10-authorship findings were similar. Group and corporate authorship were more common in general medical journals (SEE TABLE 1). Group authorship increased over time,  $P < .001$ . All papers with corporate authorship listed ( $n = 69$ ) or referenced ( $n = 2$ ) the study investigators or a writing committee. GT6, GT10, corporate and group authorship, and total author number varied significantly between journal types after controlling for year, multicenter studies, and funding,  $P < .001$  each.

**Table 1. Comparison of Authorship in General Medical and Obstetric Journals**

	General Medical Journals	Obstetric Specialty Journals	Odds Ratio (95% CI)
Research papers, No.	1,487	11,494	
Authors, mean (SD)	5.9 (3.8)	4.7 (2.3)	<sup>a</sup>
Authors, median (range)	5 (0-41)	4 (0-36)	<sup>a</sup>
Greater than 6 authors, %	34.3	13.3	3.40 (3.02-3.83)
Greater than 10 authors, %	11.7	2.3	5.57 (4.66-6.80)
Corporate authorship, %	2.8	0.3	9.11 (5.70-14.6)
Group authorship, %	10.8	2.6	4.45 (3.65-5.44)
Multicenter studies, %	12.1	6.2	2.11 (1.77-2.51)
Funded research, %	82.8	73.0	1.78 (1.55-2.05)

<sup>a</sup> $P < .0001$

**Conclusions** The proportion of multiauthored original studies regarding pregnancy is higher in general medical journals than obstetric journals, and this difference has increased over time. These findings are not accounted for by differences in funding or multicenter studies.

<sup>1</sup>Case Western Reserve University, Suite G240, Dept Ob/Gyn, MetroHealth Medical Center, 2500 MetroHealth Dr, Cleveland, OH 44109, USA, e-mail: bmm12@case.edu; <sup>2</sup>Pregnancy and Perinatology Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA; <sup>3</sup>*Obstetrics & Gynecology*, University of Utah Medical Center, Salt Lake City, UT, USA

## Bias

### The Value of Lesser-Impact-Factor Surgical Journals as a Source of Negative and Inconclusive Outcomes Reporting

Ziad Kanaan, Susan Galandiuk, Margaret Abby, Katherine Shannon, Daoud Dajani, and Hiram C. Polk

**Background** Evidence-based medicine is often used as a template for measuring the quality of medical care. Clinicians put their faith in peer-reviewed articles as quality assured and reliable knowledge. However, the peer-reviewed literature is complicated by the arduous task of clinicians to equally retrieve quality-assured positive, negative, and inconclusive reports.

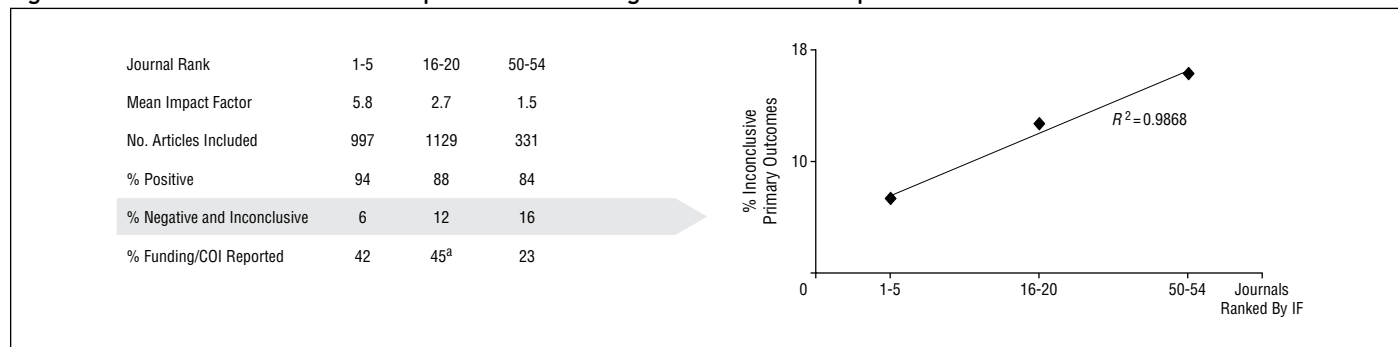
**Objectives** (1) Examine the tendency of peer-reviewed surgical journals to publish positive reports or negative and inconclusive outcome articles as a function of the journals’ impact factor (IF). (2) Examine the frequency with which surgical journal editors/publishers follow a previously published joint statement regarding funding and/or conflicts of interest (COD).

**Design** Papers from 15 surgical journals comprising 3 separate journal groups based on 2006 IF rankings were reviewed. All were published in 2007. Manuscripts were classified by 4 independent reviewers as having positive, negative, or inconclusive primary and secondary outcomes and for statements on funding/COI. Positive reports were defined as  $P < .05$ , null hypothesis rejected; negative reports defined as  $P < .05$ , null hypothesis accepted; and inconclusive reports defined as  $P > .05$ . Case reports, reviews, commentaries, and editorials were excluded. Interobserver consistency was assessed and affirmed in 10% of manuscripts.

**Results** Review of a total of 2457 articles showed an inverse correlation between impact factor and negative and inconclusive reports (FIGURE 1).

**Conclusions** This presumed bias away from opposing points of view that are essential in clinical decision making is a major weakness of current patterns of publication. This bias is further complicated by the failure of all but 1 surgical specialty journal to uniformly describe funding sources and/or COI. Lower IF-rated journals may serve a decidedly useful purpose by publishing more negative and inconclusive outcome studies. The practice of focusing disproportionately on the positive outcomes of most studies may result in unbalanced evidence.

**Figure 1. Inverse Correlation Between Impact Factors and Negative and Inclusive Reports**



<sup>a</sup>Includes the only journal surveyed with 100% funding/COI reporting.

Department of Surgery and the Price Institute of Surgical Research, University of Louisville, Louisville, KY 40292, e-mail: s0gala01@louisville.edu

<sup>1</sup>Division of Maternal-Fetal Medicine, Department of Obstetrics-Gynecology, University of California, Irvine, 101 The City Drive S, Suite 800, Building 56, Orange, CA 92868, USA, e-mail: dwing@uci.edu;  
<sup>2</sup>Obstetrics & Gynecology, Salt Lake City, UT, USA

**Differences in Editorial Board Reviewer Behavior Based on Gender**

Deborah Wing,<sup>1</sup> Rebecca Benner,<sup>2</sup> Rita Petersen,<sup>1</sup> and James Scott<sup>2</sup>

**Objective** With increasing female representation on the *Obstetrics & Gynecology's* editorial board, we evaluated whether there were differences in review behavior based on gender.

**Design** Retrospective analysis of editorial board members' reviews of original research submissions based on gender using the online management program, Editorial Manager, from January 1, 2002, through December 31, 2008. We evaluated recommendations of the editorial board members for acceptance or rejection using a 4-tier system, agreement with the editor's final decision, turnaround time from review request to submission, and editors' grades of reviews on a 5-point scale. We evaluated performance of editorial board members with advancing tenure, seeking trends in recommendations over time.

**Results** A total of 6062 manuscript reviews representing 5958 manuscripts were included; 4062 (67%) were assigned to male editorial board members and 2000 (33%) to females. There were a total of 38 editorial board members (25 men, 13 women) with tenure duration from 2 to 4.9 years, and 3 editors (2 men, 1 woman) serving 7 and 6 years, respectively. Women were less likely to accept or accept with minor revisions than were men ( $P < .003$ ). Median turnaround times were 14 (0-55) days for women and 10 (0-33) days for men ( $P < .001$ ). The editors' grades assigned to women were more often in the very good to exceptional category than for men ( $P < .0001$ ). Compared to the editors' final decisions, there was no difference based on gender with approximately 73% decision congruence overall. Men rejected more manuscripts than women with advancing tenure on the editorial board ( $P < .0001$ ).

**Conclusions** Thirteen (33%) of editorial board members for this journal are women. There are differences based on gender for the editorial board members' recommendations regarding manuscript triage, turnaround time, and editors' grades assigned. Longitudinal performance with increasing frequency of rejection recommendations with advancing tenure was found for men but not women. Overall, however, these differences do not affect the editors' ultimate decisions regarding publication of manuscripts.

**Eligibility Criteria of Randomized Controlled Trials of Acutely Ill and Hospitalized Patients With Acute Lung Injury or Sepsis**

Chris Lazongas,<sup>1,2</sup> Andrew Toren,<sup>3</sup> Ruxandra L. Pinto,<sup>1</sup> Niall D. Ferguson,<sup>2,4</sup> and Robert A. Fowler,<sup>1,2</sup>

**Objective** To evaluate the generalizability of randomized controlled trials (RCTs) of acute ill and hospitalized patients with acute lung injury or sepsis by examining RCT eligibility criteria and comparing the findings to previously established generalizability among RCTs of general medical conditions.

**Design** We searched MEDLINE for RCTs (1996-2007) involving  $\geq 50$  patients and identified 28 trials in acute lung injury and 29 trials in sepsis. Trial characteristics and eligibility criteria were abstracted. Exclusion criteria were graded as strongly, potentially, or poorly justified according to previously published guidelines.

**Results** Participants were 52.9 years old ( $\pm 13.7$ ), male (60.2%), and studied in adult intensive care units (94.7%). There were (mean, standard deviation [SD])  $12.7 \pm 6.4$  exclusion criteria per trial. Common exclusion criteria included age (86%), pregnancy or lactation (64.9%), and common medical conditions (96.5%). Shock was reason for exclusion in 24.6% of trials, weight in 21.1%. Specific medications were common exclusions among sepsis trials (62.1%) but not acute lung injury (28.6%); respiratory condition exclusion was common in acute lung injury (60.7%) but not sepsis (10.3%). Sepsis trials more commonly investigated pharmacologic agents than did those of acute lung injury (89.7% vs 39.3%) and were more commonly industry sponsored (82.8% vs 42.9%). Pharmacotherapy-based RCTs were more likely to include medication-related exclusions (odds ratio 5.87,  $P = .004$ ). Among all exclusions, 33.1% were judged strongly justified, 33.7% potentially justified, and 33.2% poorly justified. A total of 96.5% of RCTs contained more than 1 poorly justified exclusion. Compared to RCTs in the general medical literature, critical care trials exhibited fewer strongly justified exclusions (33% vs 47%), more potentially justified (33% vs 15%), and more poorly justified exclusions (37% vs 33%,  $\chi^2$  test for trend,  $P = .0066$ ).

**Conclusions** Age, gender and common medical comorbidities are common reasons for exclusion in RCTs involving the sickest of hospitalized patients. Many exclusion criteria are poorly justified. This may have important consequences for generalizability of trial results.

<sup>1</sup>Department of Critical Care Medicine, 2075 Bayview Ave, Room D-478, Sunnybrook Health Sciences Centre, Toronto, Ontario M4N 3M5, Canada, e-mail: chris.lazongas@utoronto.ca; <sup>2</sup>Department of Medicine, University of Toronto, Toronto, Ontario, Canada; <sup>3</sup>Department of Ophthalmology, University of Ottawa, Ottawa, Ontario, Canada; <sup>4</sup>Department of Medicine, Division of Respiriology, University Health Network, Toronto, Ontario, Canada.

**Multiple Publication of Positive vs Negative Trial Results in Review Articles: Influence on Apparent Weight of the Evidence**

Erick Turner

**Objective** In a previous study of selective publication, we counted trials as unpublished if they were not reported in full (stand-alone) publications. To evaluate a company's public complaints that this method was unfair, we now credit trials as published if they appear only in review publications and ask whether this mitigates the findings of selective publication for that company's drug.

**Design** Within our previously published data on antidepressant trial outcomes extracted from US Food and Drug Administration (FDA) reviews and matching full journal articles, we focused on duloxetine. We identified review articles citing the full publications using Web of Science with a cutoff date of May 2008, including only those presenting placebo-controlled efficacy outcomes. Within these articles, we determined whether the trial results were presented as positive (statistically significant) or negative (non-significant) on the primary outcome. Using Fisher exact test, we compared the proportion of positive vs negative reports according to (1) FDA reviews, (2) stand-alone publications, and (3) all (stand-alone plus review) publications.

**Results** The FDA reviewed 8 duloxetine trials and judged 4 of them positive and 4 negative. In stand-alone publications, 6 of the 8 trials were published as positive and none as negative ( $P = .085$  vs FDA tally). Within the combination of 6 stand-alone publications plus 21 review publications, outcomes from the 8 trials were reported as positive 103 times and as negative 8 times ( $P = .003$  vs FDA tally).

**Conclusions** Positive trials were fully published but negative trials were not. Results from negative trials were instead bundled with positive trials into review articles. When we counted such trials as published, we found that multiple publication of trial results significantly skewed the apparent weight of the evidence favoring drug efficacy. In the interest of fair balance, it seems reasonable to expect full publication of all trial results, regardless of trial outcome.

Department of Psychiatry and Center for Ethics in Health Care, Oregon Health & Science University, Portland, OR, USA, and Mental Health and Clinical Neurosciences Division P3MHDC, Veterans Affairs Medical Center, 3710 SW US Veterans Hospital Rd, Portland, OR 97239, USA, e-mail: turnere@ohsu.edu

**Reasons for Not Publishing Studies: A Meta-analysis of Data From Empirical Studies**

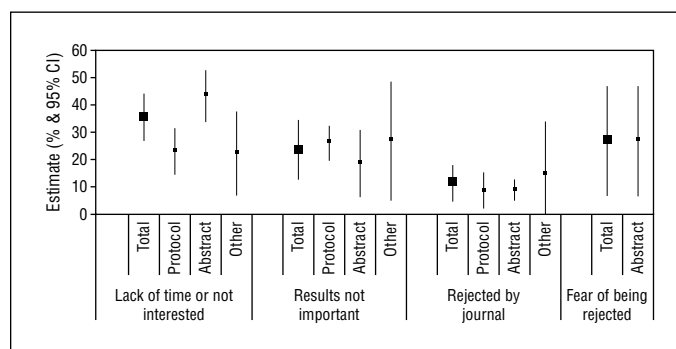
Fujian Song,<sup>1</sup> Caroline Hing,<sup>2</sup> Sheetal Parekh,<sup>3</sup> Lee Hooper,<sup>1</sup> Yoon Loke,<sup>1</sup> Jon Ryder,<sup>1</sup> Alex Sutton,<sup>4</sup> and Ian Harvey<sup>1</sup>

**Objective** To summarize data on reasons given by investigators for not publishing their studies.

**Design** As part of a comprehensive updated review of publication bias, we searched MEDLINE and the Cochrane Methodology Register Database (up to August 2008) to identify studies that provide data on reasons given by investigators for not publishing studies. References of retrieved articles were also checked for relevant studies. Percentages of specific reasons from individual studies were transformed to log odds and pooled using random-effects model.

**Results** Twenty-one studies were included (published between 1992 and 2006) including 5 studies of investigators of protocol cohorts, 11 studies of authors of meeting abstracts, and 5 studies of other or miscellaneous authors. There was significant heterogeneity in results across studies. The main reasons for nonpublication were lack of time or low priority (34.5%; 95% confidence interval [CI], 27.4%-42.3%), results not important enough (19.6%; 95% CI, 12.0%-30.4%), and journal rejection (10.2%; 95% CI, 5.5%-18.2%) (FIGURE 2). Pooled percentages of specific reasons were similar across different types of empirical studies, except that the lack of time or low interest were significantly higher in studies of meeting abstracts (43.1%; 95% CI, 35.9%-50.6%) than in studies of protocol cohorts (23.8%; 95% CI, 15.9%-34.0%) or studies of other authors (20.7%; 95% CI, 7.7%-44.9%). In the 5 studies of meeting abstracts, fear of journal rejection was given as a reason for 23.7% (95% CI, 8.9%-49.6%) of unpublished studies.

**Figure 2. Reasons for Not Publishing Studies: Pooled Results of Empirical Studies**



**Conclusions** Main reasons given by investigators for not publishing studies include lack of time or low priority and results being considered not important. Some study results remained unpublished because of journal rejection or anticipated journal rejection.

<sup>1</sup>School of Medicine, Health Policy and Practice, Faculty of Health, University of East Anglia, Norwich NR4 7TJ, UK, e-mail: fujian.song@uea.ac.uk; <sup>2</sup>Watford General Hospital, Watford, Hertfordshire UK; <sup>3</sup>School of Allied Health Professions, Faculty of Health, University of East Anglia, Norwich, UK; <sup>4</sup>Department of Health Sciences, University of Leicester, Leicester, UK

Citation Analysis

**Bibliometric Analysis of Latin-American Presence in Pediatric Publications: Geographical Distribution and Countries' Impact Factor**

Paula Otero, Norma Rossato, Pablo Duran, Fernando Ferrero, Hebe Gonzalez Pena, Susana Rodriguez, and Jose Ceriani Cernadas

**Objective** To evaluate the non-English-speaking Latin American countries' participation in pediatric journals with impact factors and that were included in MEDLINE .

**Design** All articles that were published in journals with impact factors and included in MEDLINE subset "pediatrics" between the years 1998 and 2008 were reviewed. Corresponding author or institution was used to determine the country of origin. The mean impact factor by country was calculated using data from the Journal Citation Reports database. The results obtained were adjusted for each country by population size, funds invested on research and development as percentage of gross domestic product and absolute number of researchers using the latest data available. The number of published articles was considered as an index of quantity of research productivity. The mean impact factor of the published articles was considered as an index of quality of research productivity

**Results** From 78 pediatric journals with impact factors, a total of 73,295 articles were obtained. From these articles, 1825 (2.5%) were from the 18 non-English-speaking Latin American countries. Only 7 of 18 countries had more than 20 articles published (Brazil, Argentina, Mexico, Chile, Venezuela, Colombia, and Uruguay) (TABLE 2). The country that accounted the highest number of articles was Brazil with 1055 (57.8%) followed by Argentina (17.2%) and Mexico (10.2%). The country with the highest impact factor was Chile with 2.07 followed by Uruguay (1.89), Argentina (1.62), and Brazil (1.46). When adjusted by population size, number of researchers, and research funding according to percentage of gross domestic product, Chile was the country that ranked in the highest positions for these indicators.

**Table 2. Indicators of Quantity and Quality of Research Productivity of Non-English-Speaking Latin American Countries With More Than 20 Articles Published in Pediatric Journals With Impact Factor Included in MEDLINE Between 1998 and 2008 (N = 1825)**

	Total No. of Articles (%)	Mean Impact Factor	R&D per Inhabitant (in US \$)	Articles per Population/ 1000000	Researchers per Population/ 1000000
Brazil	1055 (57.8)	1.46	65.17	5.6	1440
Argentina	314 (17.2)	1.62	52.79	7.8	1468
Mexico	187(10.2)	1.00	43.50	1.7	737
Chile	156 (8.5)	2.07	69.57	9.6	1878
Venezuela	31 (1.7)	0.97	13.95	1.2	1363
Colombia	25 (1.4)	1.16	10.77	0.6	527
Uruguay	22 (1.2)	1.89	37.02	6.4	449

R&D indicates research and development

**Conclusion** The publication rate of non-English-speaking Latin American countries in pediatric journals with impact factors is low; Brazil ranked highest in number of articles, but when other types

of analyses were done, other countries emerged as main producers of information in this discipline.

Departamento de Publicaciones, Sociedad Argentina de Pediatría, Coronel Diaz 1921, (1425) Ciudad Autonoma de Buenos Aires, Argentina, e-mail: paula.otero@hiba.org.ar

**Demographics and Intellectual Market Share of Medical Journals**

Victoria Wong<sup>1</sup> and Michael Callaham<sup>2</sup>

**Objective** Researchers and editors often need ways to identify the relative importance of journals within a discipline. Impact factor is controversial and limited in usefulness. We introduce a method of measuring a journal's relative importance within its field.

**Design** The 2007 Journal Citation Reports database was used to identify the proportion of total citations contributed by each of the medical journals within various medical categories defined by the Institute for Scientific Information (ISI). The relative number of total citations for each journal was compared to other journals within the same medical specialty. Results from the 2000 Journal Citation Reports were compared to those from 2007.

**Results** The number of journals that dominate the medical literature based on their total number of citations varies markedly among different medical subspecialties. In the field of general medicine, for example, 2 journals make up the top 42% of total citations (a pattern also common in many other disciplines). In comparison, 13 journals make up the top 42% of total citations in the field of surgery. The 2000 Journal Citation Reports data showed similar results. Citations are the "raw data" of impact on the literature but also correlate with the impact factor, a more controversial metric.

**Conclusions** We report a simple method of quantifying a journal's "intellectual market share" within disciplines. This is a useful research tool for selecting journal segments for study or educational efforts (eg, large dominant journals vs small "niche" ones). In addition, editors can measure their journals' relative importance within their field. By tracking a journal's share over time, researchers can measure a journal's citation growth and relative momentum.

<sup>1</sup>University of California, Davis, Department of Neurology, 4860 Y St, Suite 3700, Sacramento, CA 95817, USA, e-mail: vwongmd@gmail.com;

<sup>2</sup>University of California, San Francisco, San Francisco, CA, USA

**Conflicts of Interest**

**Sponsorship of Medical Textbooks by Drug or Device Companies**

Andreas Lundh and Peter Gøtzsche

**Objective** To study whether medical textbooks are sponsored by drug or device companies and, if so, whether they have tried to influence their contents.

**Design** Cross-sectional study of the medical textbooks written in Danish that are available for the pregraduate clinical courses at the University of Copenhagen and anonymous online survey of editors. For sponsored books, we also contacted the authors.

**Results** Ten of 71 medical textbooks had listed 1 or more drug or device companies as sponsors, and 1 textbook had none but was nevertheless sponsored, which we found out coincidentally. Thus, 11 books (15%) were sponsored. We contacted 11 editors, and for 8 books that had authors who were not editors, we contacted 1 author. Ten of the editors and 5 authors replied. In 2 cases, the editors had no influence on whether the book should be sponsored, as this was decided by the publisher. One of these editors was contacted 5 times by the various sponsors concerning the content of specific chapters, and in the second case the sponsor had the content of a chapter changed regarding its own drug. Two of the authors noted that they did not know that the book was sponsored. We wrote to the editors of the 60 books that did not appear to be sponsored, and 43 replied; 40 declared that there was not hidden sponsorship while 3 noted that they did not know.

**Conclusions** Sponsorship of medical textbooks is not uncommon. We regard industry sponsorship of medical textbooks as unacceptable, as it may lead to lack of academic freedom. Medical students may be particularly vulnerable to commercial influences, as they have had little or no training in commercial biases and generally believe what they read in textbooks.

Nordic Cochrane Centre, Rigshospitalet, Dept 3343, Blegdamsvej 9, DK-2100 Copenhagen, Denmark, e-mail: pcg@cochrane.dk

#### Conflict of Interest and Disclosure Policies in Psychiatry and Medicine: A Comparative Study of Peer-Reviewed Journals

Gauri Khurana,<sup>1</sup> Schuyler Henderson,<sup>2</sup> Garry Walters,<sup>3</sup> and Andres Martin<sup>4</sup>

**Objective** To characterize conflict of interest (COI) and disclosure policies published in peer-reviewed journals and to determine whether there is a qualitative difference between psychiatric and nonpsychiatric journals.

**Methods** We examined the 20 highest-ranked peer-reviewed journals in psychiatric and nonpsychiatric journals (based on 2007 impact factor). Using qualitative and quantitative approaches (including a screening instrument developed by the authors), we compared the COI and disclosure policies that appeared in print or journal Web sites through May 2009.

**Results** All journals published COI/disclosure policies that were accessible in print and online. Eight of the psychiatric journals and none of the nonpsychiatric journals required "complete" (vs "relevant") disclosure, but medical journals tended to provide more detailed information about what could constitute potential conflict and asked for broader potentially relevant funding sources. All psychiatric and 16 of the nonpsychiatric journals published COI statements for each submission. Nine psychiatric journals and 10 nonpsychiatric journals had forms for the authors; the remainder required authors to submit their own disclosures.

Three psychiatric and 8 nonpsychiatric journals specified disclosures for editors and reviewers.

**Conclusions** This preliminary study suggests that it is possible to review journals qualitatively and quantitatively to ascertain COI policies. There are variations in what information journals offer and the clarity of their expectations, and there may be field-dependent differences that affect these variations. Although COI has been a topic of substantial debate in medical fields and in larger society, there are challenges to codifying COI policies and creating standardized approaches. These challenges may reflect ongoing debates about what constitutes a COI, what needs to be disclosed, and who is responsible for disclosing COI. Further study into how journals convey COI policies and how these policies subsequently affect disclosure is warranted.

<sup>1</sup>Yale University School of Medicine, Department of Psychiatry, 300 George St, Suite 901, New Haven, CT 06511, USA, e-mail: gk109@columbia.edu; <sup>2</sup>Columbia University, Child and Adolescent Psychiatry, New York, NY, USA; <sup>3</sup>Sydney University, Discipline of Psychological Medicine, Sydney, NSW, Australia; <sup>4</sup>Yale University School of Medicine, Child Study Center, New Haven, CT, USA

#### The Impact of Disclosing Financial Ties in Research and Care: A Systematic Review

Adam Licurse,<sup>1</sup> Emma Barber,<sup>1</sup> Steven Joffe,<sup>2</sup> and Cary Gross<sup>1</sup>

**Objective** To review original, quantitative studies on the perceptions of patients, research participants, and journal readers about physicians' and investigators' financial ties (FTs) in clinical care and research.

**Design** Studies were identified by searching MEDLINE (January 1988-February 2009), Scopus, and Web of Knowledge. All English-language studies containing original, quantitative data on attitudes toward physician or investigator disclosure of FTs were included. FTs were defined as any payments made or gifts given by a company to a physician or researcher, including those directly funding research studies. We screened 6381 citations and retrieved 239 potentially eligible full articles. Of these, 18 studies met our inclusion criteria. Data were synthesized qualitatively.

**Results** Seven studies assessed patient perceptions of physician FTs. Professional gifts to physicians were viewed as unethical by 18% to 47% of patients, while 14% to 64% of patients believed FTs decrease the quality and increase the cost of care. Nine studies examined FTs in clinical research. In 5 studies, a majority of respondents believed FTs were important to disclose. Among those studies assessing willingness to participate in research, respondents were least willing to participate after a disclosure of researcher equity (18%-33% of patients). Three studies examined the impact of FTs on physicians' evaluation of research evidence. Two showed that the perceived quality of journal articles was significantly lower when FTs were disclosed ( $P < .05$ ), while in another, approximately 70% of physicians believed FTs biased clinical practice guidelines.

**Conclusions** Patients believe that FTs influence professional behavior, decrease the quality of care, and increase the cost of care. Research participants believe that FTs are important to disclose. For some, a disclosure of FTs affects their willingness to participate in research studies. Limited data suggest that FTs adversely affect physicians' assessments of the quality of journal manuscripts and of practice guidelines.

<sup>1</sup>Department of Internal Medicine, Yale Medical School, Int Med-Primary Care, PO Box 208025, New Haven, CT 06520-8025, USA, e-mail: cary.gross@yale.edu; <sup>2</sup>Department of Pediatrics, Harvard Medical School, Boston, MA, USA

e-mail: dmoher@ohri.ca; <sup>2</sup>Women's College Research Institute at Women's College Hospital, Toronto, Ontario, Canada; <sup>3</sup>Queen's University, Kingston, Ontario, Canada; <sup>4</sup>Mayo Clinic, Rochester, MN, USA; <sup>5</sup>Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; <sup>6</sup>School of Health Policy and Management, York University, Toronto, Ontario, Canada; <sup>7</sup>Emergency Department, University Health Network, Toronto, Ontario, Canada; <sup>8</sup>Department of Family and Community Medicine, University of Toronto, Toronto, Ontario, Canada; <sup>9</sup>Toronto Rehabilitation Institute, Toronto, Ontario, Canada; <sup>10</sup>Joint Centre for Bioethics, University of Toronto, Toronto, Ontario, Canada; <sup>11</sup>Department of Physiology University of Toronto, Toronto, Ontario, Canada; <sup>12</sup>Baycrest Kunin-Lunenfeld Applied Research Unit, Toronto, Ontario, Canada

### Screening Investigator Financial Conflict of Interest: A Checklist for Authors, Editors, and Readers

David Moher,<sup>1</sup> Paula Rochon,<sup>2</sup> John Hoey,<sup>3</sup> An-Wen Chan,<sup>4</sup> Lorraine Ferris,<sup>5</sup> Joel Lexchin,<sup>6-8</sup> Marleen Van Laethem,<sup>9,10</sup> Sunila Kalkar,<sup>2</sup> Melanie Sekeres,<sup>11</sup> Wei Wu,<sup>2</sup> and Andrea Gruneir<sup>2,12</sup>

**Objective** To develop a simple, comprehensive tool to help investigators identify and to report the extent and types of financial conflicts of interest (fCOI) present in current funded research.

**Design** Between January 2007 and April 2009, we developed the fCOI Checklist using a 3-phase process (premeeting item generation, consensus meeting, and postmeeting consolidation). The checklist items were initially generated by our research team based primarily on published literature of initiatives that targeted specific aspects of fCOI. When required items were not available from these sources, we created the item. We used a modified Delphi process from team members and invited external panel members to revise the checklist. The reviewers used a 5-point adjectival rating scale (1, least important, to 5, most important) and also provided free text suggestions to improve the item for 2 sequential checklist iterations. Twenty-eight people including representation from journals, law, ethics, and research integrity participated in the consensus meeting. In the postmeeting phase, the revised checklist was piloted for usability, and a Web-based version of checklist was created. In the final meeting the revised fCOI Checklist was finalized for investigators conducting clinical trials.

**Results** The final fCOI checklist is to be completed by an investigator for an individual study. It contains 4 sections (ie, administrative, study, personal financial, and authorship information). These sections are divided into 6 modules containing 14 items and their related subitems. Different modules within the fCOI checklist should be completed at different transition points over the course of the study and updated information to be appended to the originally completed fCOI checklist. The checklist can be completed in fewer than 20 minutes.

**Conclusions** We consider this fCOI Checklist to be a living document. We invite comments and suggestions to improve the checklist and suggestions for adaptations.

<sup>1</sup>Clinical Epidemiology Methods Centre, Ottawa Health Research Institute, The Ottawa Hospital, General Campus, Critical Care Wing (Eye Institute), 6th Floor, 501 Smyth Rd, Ottawa, Ontario K1H 8L6, Canada,

### Data Sharing

#### Medical Journal Editor Perspectives on Sharing Results Data From Patient-Oriented Research: A WAME Survey

Karmela Krleža-Jerić,<sup>1</sup> Ida Sim,<sup>2</sup> Ana Marušić,<sup>3</sup> Ludovic Reveiz,<sup>4</sup> and Carlos Granados<sup>5</sup>

**Objective** This study is assessing perceptions, policies, and practices of medical journal editors regarding various levels of public disclosure of patient-oriented research data.

**Design** We surveyed editors of member journals of the World Association of Medical Editors (WAME). The 30-item online survey inquired about journal characteristics and editors' views on the public posting of study results datasets including issues of timing, ideal formats, and potential dangers. Initially, 461 WAME members were recruited starting in March 2009 via the WAME listserv. Four reminders followed: 2 via listserv and 2 directly using e-mails extracted from journals' Web sites. We supplemented questionnaire data with journal characteristics gathered from the Web of Knowledge, PubMed, and SHERPA/RoMEO. We used descriptive statistics to examine journals' and editors' characteristics and frequencies of reported perceptions, policies, and practices.

**Results** As of June 9, 2009, the survey is ongoing. From the original 461 journals, we excluded 131 due to absent or incorrect e-mail addresses. We have now received 102 responses from 89 different journals and 29 countries. Most responders were active editors in chief (47%) or active editors (36%). Most journals are in English (89%) and are indexed in MEDLINE (57%), and 48% had full open access. Thirty-four percent of journals that publish studies involving human participants require trials to be registered prior to inception, while 19% require registration even if they are registered retrospectively in a trial registry approved by the International Committee of Medical Journal Editors or the World Health Organization. Only 7/89 (8%) and 2/89 (2%) journals currently require summary level data and participant level data, respectively. Only 19% of journals require authors to specify their data-sharing plan. We extracted journal characteristics from citation databases for 105 randomly selected WAME member journals. Eighty-six percent of these journals have online publication, 44% have full open access; median impact factor is 2.75 among journals that reported such data.

**Conclusions** This survey is limited by the moderate response rate, which is due to uncertainty about the actual numbers of active members and journals on the listserv and also whether they publish clinical research. Nevertheless, the findings of this survey may be used to develop results reporting policies by WAME and other organizations and to inform the development of international standards for public disclosure of trial results.

<sup>1</sup>Canadian Institutes of Health Research, 160 Elgin St, Ottawa, Ontario K1A 0W9, Canada, e-mail: karmela.krleza-jeric@cihr-irsc.gc.ca; <sup>2</sup>University of California San Francisco, San Francisco, CA, USA; <sup>3</sup>*Croatian Medical Journal* and Croatian Centre for Global Health, School of Medicine, University of Split, Split, Croatia; <sup>4</sup>Cochrane Collaboration Branch, Research Institute, Sanitas University Foundation, Bogota, Colombia; <sup>5</sup>Research Institute, National University of Colombia, Bogota, Colombia

## Editorial Decision Process

### Consistency of Decision Making by Editors: The Relation Between Reviewers' and Editors' Ratings and Future (10 Years) Citation

Tobias Opthof<sup>1,2</sup> and Ruben Coronel<sup>1</sup>

**Objective** To test the consistency of editorial decisions and to assess the relation between reviewers' and editors' ratings and citations.

**Design** The editors of *Cardiovascular Research* performed an analysis of reviewers' and editors' ratings of 169 original manuscripts consecutively submitted between October and December 1997. First, each editor (7) rated these manuscripts, leading to a combined editors' score (range, 0%-100%). Next, reviewers' reports (3) led to a reviewers' priority rating from 0% (3 low), 33% (1 low, 2 high), 67% (2 high, 1 low), or 100% (3 high). All ratings were compared and associated with citations obtained during 10 years. An editor's decisions (by 1 specific editor) on 21 selected, nonrejected, manuscripts were redone in a blinded manner by all editors 2 months later (without consequence for the manuscripts), including the editor who had made the original decision.

**Results** From 169 manuscripts, 56 (33%) were published in 1998-1999 (53/3). The same editor who had previously decided to accept the manuscript decided to reject it 33% of the time when he considered the manuscript again based on the same materials (67% for the team decision by majority vote). The editor's ratings had a very weak relation with the reviewer's ratings. Neither the reviewer's ratings nor the editor's ratings were significantly correlated with 10 years' citations. Only manuscripts with both an editor's ratings >50% and reviewer's ratings ≥67% were cited more than the other manuscripts (29.5 ± 5.5, n = 30, vs 15.0 ± 1.9, n = 26;  $P < .025$ ).

**Conclusions** Individual editors' decisions are far from consistent, team decisions comply poorly with individual editors' decisions, editors' ratings do not predict reviewers' ratings, neither reviewers' ratings nor editors' ratings predict citation, and combined reviewers' and editors' ratings poorly predict future citation.

<sup>1</sup>Department of Experimental Cardiology, Center for Heart Failure Research, Academic Medical Center, Meibergdreef 9, Room K2-105, 1105 AZ Amsterdam, the Netherlands, e-mail: t.opthof@inter.nl.net; <sup>2</sup>Department of Medical Physiology, University Medical Center, Utrecht, the Netherlands

### Evaluation of Editors' Judgment on Quality of Articles

Heidi Logothetti,<sup>1</sup> Sheryl Martin,<sup>1</sup> Rebecca Benner,<sup>1</sup> James Scott,<sup>1</sup> John Queenan,<sup>1</sup> and Catherine Spong<sup>2</sup>

**Objective** To evaluate whether the best articles as judged by the editors of a peer-reviewed medical specialty journal (*Obstetrics & Gynecology*) correspond with the best articles as defined by the highest number of citations and online accesses.

**Design** Every year, the editors of *Obstetrics & Gynecology* nominate 12 to 15 original research articles in the journal that originate from a US-based institution for an award (the Pitkin Award). Nominees are selected for their scientific merit, importance to the specialty, study design and methodology, presentation of results, soundness of conclusions, and writing style. An independent committee selects the 4 to 5 best papers from this group. In this case-control study, nominated articles from 2002 to 2007 were matched with the next consecutive journal article eligible for the award but not nominated (controls). The number of citations and online accesses was compared between awardees and controls, and between nominees (including awardees) and controls, using Wilcoxon signed-rank tests. Citation data were obtained from the Web of Knowledge Journal Citation Reports.

**Results** The 25 award-winning articles published between 2002 and 2007 received more citations when compared to the matched controls (median [range], 17 [0-137], compared to 7 [0-52],  $P = .02$ ). The online accesses were similar between the 2 groups (4675 [1655-10226] compared to 4368 [888-11621],  $P = .31$ , respectively). The 119 nominees received more citations (15 [0-170] compared to 11 [0-62],  $P < .001$ ) and online accesses (4737 [0-15776] compared to 3671 [0-14042],  $P < .001$ ) than their matched controls.

**Conclusions** Articles subjectively judged by the editors of *Obstetrics & Gynecology* to be the best performed well by 2 objective standards, indicating that the editors accurately assessed the quality of manuscripts. Such ability is critical to journal editors, who decide which studies are published and disseminated.

<sup>1</sup>*Obstetrics & Gynecology*, 409 12th St SW, Washington, DC 20024-2188, USA, e-mail: hlogothetti@greenjournal.org; <sup>2</sup>Pregnancy and Perinatology Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA

### Subjectivity and the Science of Peer Review: What Qualitative Research Can Tell Us About Editorial Processes at Major Medical Journals

Wendy Lipworth

**Objective** Editorial and peer review of manuscripts has recently become a popular and important subject of academic research. While some qualitative research has been conducted, most of this research is quantitative. Many important insights have been derived from quantitative research, but qualitative methods may be better suited to understanding the nuances and complexities of open systems such as manuscript review.

**Design** Qualitative methods, based on grounded theory, were used to carry out an in-depth and inductive analysis of the manuscript review process. Data sources consisted of (1) in-depth, open-ended interviews with journal editors, peer reviewers, and authors from a range of international journals and (2) written peer review documents and editorial deliberations from the *Lancet* (chosen on the grounds that it is a “model” journal). This combination of sources was felt to provide a rich and detailed insight into the manuscript review process across medical publishing.

**Results** Despite efforts to ensure that manuscript review is scientific, the review process is characterized by (1) complex negotiations of epistemic authority; (2) dynamic, shifting, and contextually specific relationships of power and vulnerability; (3) reciprocal moral responsibilities; and (4) judgments that are unavoidably both prejudiced (in the neutral sense of the term) and strongly intuitive.

**Conclusions** Qualitative research into manuscript review is an important adjunct to quantitative studies. The results of this qualitative study have the potential both to challenge existing editorial assumptions, attitudes, and practices and to inform further quantitative research into the manuscript review process.

Centre for Values, Ethics and the Law in Medicine, Medical Foundation Building (K25), University of Sydney, Sydney NSW 2006, Australia, e-mail: w.lipworth@usyd.edu.au

## Ethical Concerns

### Informed Consent in Clinical Studies Published in the *Chinese Medical Journal*: Comparison of 2001-2004 and 2005-2008

Mouyue WANG

**Objective** To investigate and compare the informed consent in published clinical studies in the *Chinese Medical Journal* from 2001-2004 to 2005-2008.

**Design** The *Chinese Medical Journal* is a top medical journal in China published in English by the Chinese Medical Association. This survey aimed to evaluate the status and changes of reporting informed consent in the journal from 2001-2004 to 2005-2008. All full-text (with abstracts) original articles on clinical studies related to diseases' diagnosis, treatment, and prognosis published in the journal during these periods were included. Animal and in vitro studies as well as non-full text articles were excluded. Abstract, introduction, and methods sections were reviewed twice for each paper.

**Results** A total of 235 papers published in 2001-2004 were investigated, among which 34 reported having obtained informed

consent from participants or their guardians, accounting for 14.5%; 30.7% (92/300) of papers published in 2005-2008 reported having obtained informed consent; significant improvement in reporting informed consent was found during these periods ( $\chi^2 = 12.13$ ,  $P < .001$ ). The proportion of reporting informed consent in papers of prospective design during 2001-2004 was 43.5% (27/62), in retrospective design was 4.0% (7/173), significant difference was found ( $\chi^2 = 37.73$ ,  $P < .001$ ); the proportions during 2005-2008 were 67.1% (51/76) and 18.3% (41/224), respectively, significant difference was also found ( $\chi^2 = 29.13$ ,  $P < .001$ ). The overall proportion of reporting informed consent in prospective studies during 2001-2008 was 56.5% (78/138), significantly higher than that in retrospective ones (12.1%, 48/397,  $\chi^2 = 60.45$ ,  $P < .001$ ).

**Conclusions** The proportions of reporting informed consent in published clinical studies in the *Chinese Medical Journal* were both low in 2001-2004 and 2005-2008, though significant improvement was found in 2005-2008 compared to 2001-2004. The proportions in prospective studies were significantly higher than that in retrospective ones.

*Chinese Medical Journal*, Chinese Medical Association, 42 Dongsu Xidajie, Beijing 100710, China, e-mail: cmawmy@sohu.com; cmawmy@cma.org.cn

### Ethical Concerns of Nursing Journal Reviewers: An International Survey

Marion E. Broome,<sup>1</sup> Molly Dougherty,<sup>2</sup> Margaret Kearney,<sup>3</sup> Margaret Freda,<sup>4</sup> and Judith Baggs<sup>5</sup>

**Objective** Editors of scientific literature rely heavily on peer reviewers to evaluate the integrity of research conduct and validity of findings in manuscript submissions. The purpose of this study was to describe the ethical concerns of reviewers for nursing journals.

**Design** This descriptive cross-sectional study was an anonymous online survey. The findings reported here were part of a larger investigation of experiences of reviewers. Fifty-two editors of nursing journals (6 outside the United States) agreed to invite their review panels to participate. A 69-item forced-choice and open-ended item survey developed by the authors based on the literature was pilot tested with 18 reviewers before being entered into an online survey program. A total of 1675 reviewers responded with useable surveys. Ninety-one percent of the respondents were women, and 74% from the United States; the remaining 26% represented 44 different countries. Six questions elicited responses about ethical issues, such as conflict of interest, protection of human subjects, plagiarism, duplicate publication, and misrepresentation of data. Reviewers indicated whether they had experienced such a concern and notified the editor, how satisfied they were with the outcome, and provided specific examples.

**Results** TABLE 3 presents the findings related to the concerns identified and their outcomes. Approximately 20% of reviewers had experience with various ethical dilemmas. Although the majority reported their concerns to the editor, not all did, and not all were satisfied with the outcomes. The most commonly reported concern was perceived inadequate protection of human subjects. The least common was plagiarism, but this concern was the one most

often reported to the editor and least often led to a satisfactory outcome. Qualitative responses at the end of the survey indicate this lack of satisfaction was most commonly related feedback provided on the resolution by the editor.

**Table 3. Ethical Issues Identified by Reviewers and Satisfaction With Outcome of Notification**

Ethical Concern	No. Identified Ethical Issue	Notified Editor, %	Satisfied With Outcome, %
Conflict of interest	331 (22.5%)	92	94
Human subjects	346 (24%)	96	82
Duplicate publication	305 (21%)	94	90
Plagiarism	232 (16%)	98	80
Data representation	252 (18%)	94	85

**Conclusion** The findings from this study suggest several areas editors should take note, including follow-up with reviewers when they identify ethical concerns about a manuscript.

<sup>1</sup>Indiana University School of Nursing, 1111 Middle Dr, Indianapolis, IN 46202, USA, e-mail: mbroome@iupui.edu; <sup>2</sup>University of North Carolina at Chapel Hill, School of Nursing, Hillsborough, NC, USA; <sup>3</sup>University of Rochester School of Nursing, Rochester, NY, USA; <sup>4</sup>Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY, USA; <sup>5</sup>Oregon Health Sciences University School of Nursing, Portland, OR, USA

**Must Research Using Publicly Available Anonymous Databases Undergo IRB Review? Views From Journals and IRBs**

Ingrid Nygaard and Sheryl Martin

**Objective** According to US federal regulations, research involving publicly available data is exempt from institutional review board (IRB) oversight. Interpretation of this regulation seems inconsistent. Our aim is to describe journal and IRB policies related to whether IRB review is required for such studies.

**Design** We evaluated all original contributions using publicly available databases published between July 1, 2008, and December 31, 2008, in 6 high-impact, patient-oriented research journals read by obstetrician-gynecologists to determine whether methods sections specifically described IRB exemption or approval. We excluded meta-analyses, decision analyses, databases requiring approval for use, and public health surveillance summary data. We accessed Web sites for all IRBs associated with the 126 accredited MD-granting US medical schools to describe local policies. We examined Web sites for top 10 obstetric/gynecologic journals according to impact factor and 5 high-impact general journals read by obstetrician-gynecologists to determine journals' IRB policies.

**Results** Of 447 original research studies, 19 met inclusion criteria. Authors noted that 11 were reviewed by the IRB and 8 were not. Of those reviewed, 7 were considered exempt. Of IRB Web sites, 93 (74%) stated that only the IRB could determine exemption, and 28 (22%) gave investigators or departmental leaders this authority. We were unable to access 5 Web sites. Four journal Web sites clearly stated that authors must document in all manuscripts formal IRB exemption or approval, and 1 required a separate

statement of same. Authors were instructed in 4 journals to document IRB status in manuscripts "if applicable" and in 2 journals that ethical consent was expected. Four journals did not mention IRB in instructions.

**Conclusions** US federal regulations do not specify which body is responsible for determining research exemption. Institutional review board policies are not uniform and continue to be in flux. Journals should strive for improved consistency and transparency in their requirements of IRB oversight for research using publicly available databases.

University of Utah School of Medicine, Obstetrics & Gynecology, 30 N 1900 E, Salt Lake City, UT 84108, USA, e-mail: ingrid.nygaard@hsc.utah.edu

**Exposing a Loophole: Ethics Committee Review Requirements for Published Abstracts of Human Research Presented at Major Medical Association Meetings**

Rachel N. Simmons<sup>1</sup> and Robert P. Dellavalle<sup>2,3</sup>

**Objective** To examine ethics committee (eg, institutional review board [IRB]) review requirements for human research presented at major medical society meetings and publication of meeting abstracts reporting human research in affiliated medical journals.

**Design** For this descriptive study, we compiled a list of the 100 medical journals with the highest measures of SCImago Journal Rank (SJR). From this list, we identified journals affiliated with medical societies. The sample represented a variety of specialties, the majority being widely recognized US medical societies. Between January 5, 2009, and February 20, 2009, the Web sites of each medical journal and affiliated society were accessed and the Information for Authors or Abstract Submission Guidelines sections were digitally saved and examined. If information on the Web site was unavailable, the authors contacted the journal or society to clarify ethics committee approval requirements.

**Results** Ethics committee approval requirements for human research submitted to academic journals exceeded those of abstracts submitted to affiliated medical society meetings (100% [27/27] vs 37% [10/27] required approval). Twelve journals or their supplements (44%) published abstracts of research presented at the medical society meetings; ethics committee approval was not required by a majority of medical societies prior to the publication of meeting abstracts describing human research in affiliated journals (58% [7/12]). None of the journals or societies in the study required documentation of ethics committee approval prior to publication (TABLE 4).

**Conclusions** Although all of the medical journals in the study required ethics committee approval for manuscripts describing research on human subjects, a loophole exists regarding abstracts from meetings. Many academic societies did not explicitly require ethics committee approval for human research presentations at society meetings and some of these abstracts were published in affiliated journals. Ethical committee review should be confirmed for all human research presentations at medical meetings and resulting published abstracts.

**Table 4. Journals Publishing Meeting Abstracts of Human Research Without Requiring Ethics Committee Review**

Medical Societies Not Requiring Ethics Committee Approval for Abstracts	Country	Affiliated Journals (That Later Publish Meeting Abstracts)	2007 ISI Impact Factor for Journal
European Association for the Study of the Liver	Switzerland	<i>Hepatology</i>	10.7
American Thoracic Society	United States	<i>American Journal of Respiratory and Critical Care Medicine</i>	9.1
American Diabetes Association	United States	<i>Diabetes</i> <sup>a</sup>	8.3
American Heart Association	United States	<i>Arteriosclerosis, Thrombosis, and Vascular Biology</i> <sup>a</sup>	7.2
American Society of Nephrology	United States	<i>Journal of the American Society of Nephrology</i> <sup>a</sup>	7.1
Society of Critical Care Medicine	United States	<i>Critical Care Medicine</i> <sup>a</sup>	6.3
International Society for Biological Therapy of Cancer	United States	<i>Journal of Immunotherapy</i>	4.8

<sup>a</sup>Meeting abstracts are printed in a supplement to the journal.

<sup>1</sup>University of Florida College of Medicine, Gainesville, FL, USA;

<sup>2</sup>Department of Veterans Affairs Medical Center, Denver, CO, USA;

<sup>3</sup>Department of Dermatology, University of Colorado Denver, School of Medicine, PO Box 6510, Mail Stop F703, Aurora, CO 80045-0510, USA, e-mail: robert.dellavalle@uchsc.edu

## Funding/Grant Peer Review

### Quality Assurance of Grant Applications: Confirmation of Publication Records

Liza Chan,<sup>1,2</sup> Kathryn Graham,<sup>1</sup> Pam Valentine,<sup>1</sup> and Jacques Magnan<sup>1</sup>

**Objective** There is a scarcity of literature on the integrity of publication records submitted in grant applications. This study aimed to determine the prevalence and extent of observed discrepancies in the publication lists of grant applications at the Alberta Heritage Foundation of Medical Research (AHFMR), a provincial health research funding agency.

**Design** A retrospective review of all applications in the 2007 AHFMR investigator awards competition was conducted. Applicants' self-reported peer-reviewed publications were examined against PubMed data following an in-house protocol. The entire publication list was checked for all applications in the junior award categories. For the senior award applications, which often had extensive publication listings, the most recent 20 publications or publications from the last 5 years (whichever had the smaller number) were checked. Types and number of discrepancies (eg, authorship, titles) were documented, classified, and analyzed.

**Results** Among the 125 applications, 29 (23%) had no discrepancies identified. A total of 376 discrepancies were detected within the 1928 publications checked. The most common discrepancies detected were (1) different article titles (177/376, 47%), (2) omis-

sion of coauthors (92/376, 24%), and (3) changed authorship order (69/376, 18%). There were no nonexistent publications nor false claims of authorship observed.

**Conclusions** This study reveals the types and prevalence of discrepancies in self-reported publication records among grant applications at AHFMR in 2007. The results of this study have led to the development of a quality assurance protocol within AHFMR that includes the addition of an integrity statement in the application form requiring signatures of the applicant and the sponsoring institutions. Data gathering has started in subsequent competition years for comparison. Funding agencies are challenged to implement quality assurance systems to verify the integrity of information collected and ensure that these systems are both feasible and cost effective.

<sup>1</sup>Heritage Foundation for Medical Research, 1500, 10104-103 Ave NW, Edmonton, Alberta T5J 4A7, Canada, e-mail: liza.chan@ahfmr.ab.ca;

<sup>2</sup>John W. Scott Health Sciences Library, University of Alberta, Edmonton, Alberta, Canada

## Instructions for Authors

### Completeness of Journal Instructions for Authors

Jane C. Wiggs

**Objective** Authors, authors' editors, and manuscript preparers rely on journals' Instructions for Authors to be current and complete. Most Instructions for Authors include details of manuscript submission and format. The purpose of this study was to determine whether high-quality (high-impact-factor [IF]) journals have high-quality (comprehensive) Instructions for Authors, with content beyond submission and format guidance.

**Design** Ten criteria were sought in Instructions for Authors of the highest-IF clinical journals publishing original research according to the 2007 Journal Citation Reports (JCR). The qualifying journal with the highest IF in each clinically oriented category was included. Criteria sought were reference in the IFA to (1) the journal's peer review process, (2) clinical trial registration, (3) open access (free or paid), (4) public access (in response to the National Institutes of Health mandate), (5) the sponsor's role in the study, (6) authors' access to data, (7) Web site of the International Committee of Medical Journal Editors (ICMJE) for policy (not style) matters, (8) figure integrity or image manipulation, (9) racial or sex bias in subject selection, and (10) detailed author contributions for publication. The main outcome measures were the mean number and the frequency of criteria met.

**Results** Included in the study were 46 journals from 51 JCR categories (4 journals had the highest IF in more than 1 category). Impact factors ranged from 2.217 to 52.589. The mean number of criteria met was 3.9 (median, 4; mode, 4; range, 1-9). Criteria were met as follows: the journal's peer-review process (n = 37), public access mandate (n = 31), open access (n = 25), clinical trial registration (n = 22), ICMJE Web site for policy (n = 19), figure integrity or image manipulation (n = 13), detailed author contributions for publication (n = 11), sponsor's role (n = 10), authors' access to data (n = 7), and racial or sex bias (n = 4).

**Conclusion** Many journals, irrespective of quality, publish Instructions for Authors that lack important information.

Mayo Clinic, Section of Scientific Publications, 4500 San Pablo Rd, Jacksonville, FL 32224, USA, e-mail: wiggs.jane@mayo.edu

**High-Impact-Factor Journals Offer Limited Guidance to Authors Reporting Survey Research**

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

**Objective** Reporting guidelines, defined as a checklist, flow diagram, or explicit text to guide authors who are reporting a specific type of research using explicit methodology, have been developed to inform reporting for a variety of study designs. As part of a broader review of the literature to identify whether reporting guidelines exist for survey research, our aim was to examine the extent to which leading medical journals provide guidance for reporting survey research.

**Design** We examined Instructions to Authors Web pages between January 12 and February 9, 2009, for the top 5 journals (by impact factor) from 33 medical specialties identified through Web of Knowledge. All text containing the search terms survey, questionnaire, response rate, and nonresponder was extracted. Web pages were also hand-searched for reference to reporting guidelines for any study design. Additionally, we used PubMed to verify whether or not the journals publish survey research articles.

**Results** Of 165 high-impact journals identified, 83% (137/165) publish survey research articles. Ten percent (17/165) of the Instructions to Authors Web pages contained 1 or more of the search terms. Four percent (7/165) contained 1 or more search term(s) only; that is, no guidance or directive(s). Another 4% (7/165) contained 1 brief statement, directive, or reference(s) relevant to reporting survey research. A further 2% (3/165) contained more than 1 directive relevant to reporting survey research. While many (95/165) of the journals reference at least 1 reporting guideline for other types of study designs, none refer to a reporting guideline for survey research.

**Conclusions** The majority of high-impact-factor journals publish survey research. Most Instructions to Authors for these journals do not offer guidance nor refer to guidelines for reporting survey research.

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

**Open Access**

**Impact of Free Access of Articles on Content and Impact Factor of an Indian Biomedical Journal**

Anju Sharma and K. Satyanarayana

**Objective** To assess the impact of free access of articles on the content and impact factor (IF) of the *Indian Journal of Medical Research (IJMR)*.

**Design** The study period was divided into 2 periods: pre-open access (2000-2003) and post-open access (2004-2008). The parameters compared were IF, articles submitted, published year, international contribution, and subscription/royalty. During the study the standard, procedures for submission and peer review remained the same.

**Results** There was an increase in the number of manuscripts submitted post-2004 when the full text of the journal was made available for free online. The increase in submission (base 2000) was 94% to 182% in 2005-2008. Similarly, international contributions rose from none in 2000 to 13% in 2005 and 28% in 2008. The reviewers' base also became strong and international. Up to 2003 the reviewers' base was largely Indian, and in 2008, 58% reviewers were from countries other than India. The IF of the *IJMR* gradually increased from 0.4 (2003) to 1.67 in 2008 (TABLE 5). Subscription/royalty remained more or less same during the 2 study periods.

**Table 5. Impact of Free Access of Content on the Overall Improvement of the *Indian Journal of Medical Research***

Year	Impact Factor	International Contributions Published Papers, % of total	% Increase in Submissions, 2000 Base Year	% Increase in Published Articles, 2000 Base Year	International Reviewers, % of Total
2000	0.38	0			0
2001	0.34	0	10	0	0
2002	0.45	0	22	0	0
2003	0.45	8	49	0	0
2004	0.60	4	61	5	5.6
2005	0.87	13	94	97	24
2006	1.22	23	115	204	40
2007	1.67	24	142	188	37
2008	Awaited	28	182	274	58

**Conclusion** Making the full text of the journal available for free online appeared to have contributed significantly in improving the quality, content, and outreach of the *IJMR*, as the review policies and procedures remained the same.

Division of Publication & Information, *Indian Journal of Medical Research*, Indian Council of Medical Research, V. Ramalingaswami Bhawan, Ansari Nagar, New Delhi 110 029, India, e-mail: anjusharma2@gmail.com

**Peer Review**

**ET Study (Enhancing Transparency of Biomedical Journals)**

Erik Cobo,<sup>1,2</sup> Agustí Urrutia,<sup>2</sup> Albert Selva-O'Callaghan,<sup>2</sup> Francesc Cardellach,<sup>2</sup> Josep Maria Ribera,<sup>2</sup> Jordi Cortés,<sup>1</sup> Francesc Miras,<sup>1</sup> Celestino Rey-Joly,<sup>2</sup> and Miquel Vilardell<sup>2</sup>

**Objective** The aim of reporting guidelines (RG) is to enhance the quality and transparency of health research, which may be achieved by including a senior statistician who asks authors to provide information about incomplete or missing RG items. The objective of this study was to investigate the effect of additional review using RG checklists on the quality of manuscripts published in a weekly medical journal with 1.3 impact factor and no specific requirements to follow RG.

**Design** A masked randomized trial of original research manuscripts conditionally accepted for publication in *Medicina Clinica* after standard peer review. Half the manuscripts received an extra review performed by a senior statistician. The primary end point was a change from initial to final Goodman quality rating as rated by 3 statisticians who were masked to the randomized group. Allocation concealment: the editorial committee decision was made after peer review was performed and before randomization without information about the additional RG review. Later decisions were aware of the additional RG review only in the treated papers. Random allocation: minimization of differences in initial overall Goodman quality and study type (4 groups: intervention, longitudinal, transversal, and other). The sample size calculation indicated that 50 papers per group allowed 80% power to detect a difference of means equivalent to 55% of the change standard deviation.

**Results** From May 2008 to April 2009, 126 consecutive papers included the extra RG review. From them, 34 were rejected on the basis of the conventional review and 92 were randomized (for 2 papers the final version was not sent within the scheduled time). Among the extra review and conventional groups, 44.9% (22/49) and 19.5% (8/41) of papers, respectively, were improved from baseline in Goodman overall quality score (OR 3.32, 95% confidence interval [CI], 1.19 - 10.07), but the main analysis failed to show a significant difference on the means (0.26, 95% CI, -0.08 - 0.61, SD=1). (FIGURE 3.)

**Conclusion** There is some evidence that the extra review improves paper quality, although its effect size seems to be moderate and smaller than hypothesized.

<sup>1</sup>Department of Statistics and Operations Research, UPC, Jordi Girona 1-5 (C5), 08034 Barcelona, Spain, e-mail: erik.cobo@upc.edu; <sup>2</sup>Editorial Committee, *Medicina Clinica*, Barcelona, Spain

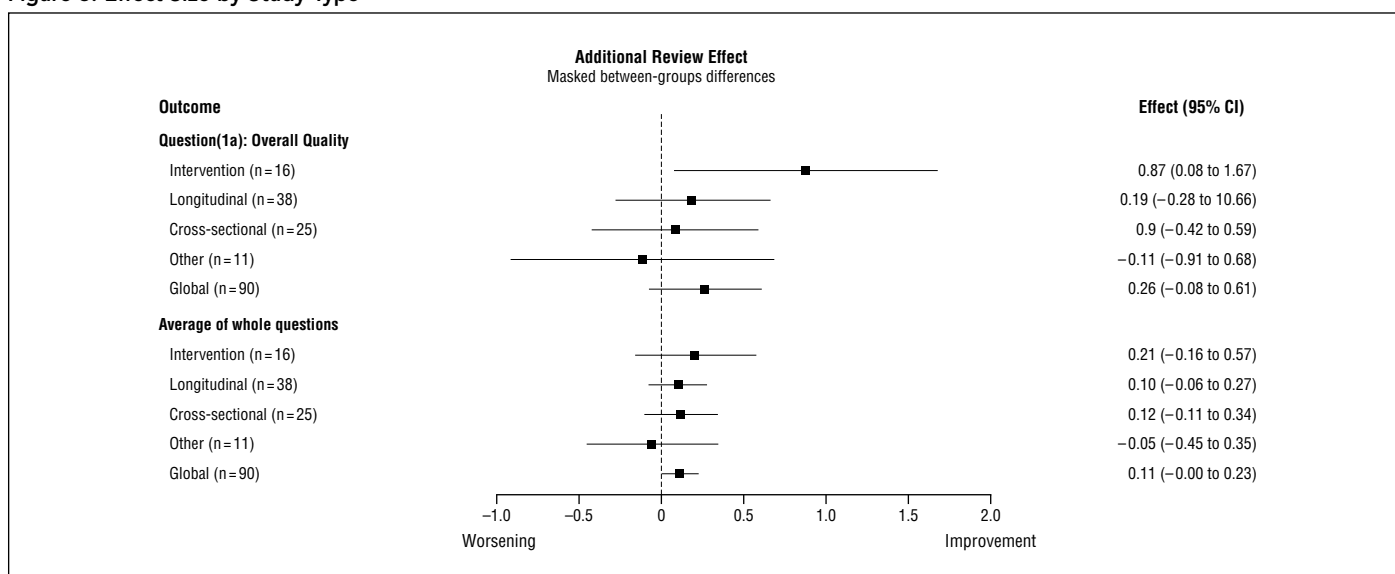
**Diversity and Quality of Reviews for a Generalist Journal**

William Phillips,<sup>1</sup> Robin Gotler,<sup>2</sup> Laura McLellan,<sup>2</sup> and Stephen Zyzanski<sup>2</sup>

**Objective** General medical journals must evaluate research with a wide variety of topics and methods. Diverse expertise and viewpoints can be valuable, but it is unknown if individuals from other fields can perform quality reviews. We compared editor assessments of reviews done by reviewers in diverse fields with those done by experts in the journal's core field.

**Design** Retrospective study of all reviews done for *Annals of Family Medicine* in 2006-2008. Editors contemporaneously made global subjective assessments of each review on a 5-point Likert-type scale: 5 = excellent, 4 = good, 3 = average, 2 = poor, 1 = unacceptable. Reviewers identified themselves in 1 primary field. We grouped fields by how closely they were related to the core audience of the journal, family medicine/general practice (FP/GP). Nonmedical reviewers included fields such as anthropology, business, communications, economics, education, English, ethics, informatics, law, library science, and social sciences. Other reviewer groups comprised the following categories: Other Primary Care Physicians (pediatrics, internal medicine), Physician Specialties, Non-Physician Health Professionals, and Public Health Sciences. Our primary outcome was proportion of reviews rated good or excellent (4 or 5). We tested differences with  $\chi^2$  tests for proportions and t test for means.

Figure 3. Effect Size by Study Type



Forest plot of the extra review effect size (mean difference between treated and control groups) on the overall quality item (1 to 5) in Goodman scale and on the average of all valid Goodman questions.

**Results** Over 3 years 767 reviewers returned 981 reviews on 563 manuscripts. For the 959 reviews rated by editors, 70% (672/959) were rated good/excellent (95% confidence interval [CI], 0.67-0.73), mean, 3.91 (95% CI, 3.84-3.97), standard deviation (SD) 1.05, median, 4, range, 1-5. FP/GP reviewers returned 70% good/excellent reviews (95% CI, 0.66-0.73), mean 3.91. Nonmedical reviewers returned 70% (46/66) good/excellent reviews (95% CI, 0.58-0.80), mean, 3.94 ( $\chi^2 P = .65$ ,  $t$  test  $P = .81$ ). Other groups returned reviews of very similar quality (TABLE 6).

**Table 6. Review Ratings by Reviewer Group**

Reviewer Group	Good/Excellent Reviews (95% CI) <sup>a</sup>	Total Reviews	Mean (95% CI)	SD	Median	Range
All Reviewers	70% (672) (0.67-0.73)	959	3.91 (3.84-3.97)	1.05	4	1,5
Family Medicine <sup>b</sup>	70% (430) (0.66-0.73)	618	3.91 (3.83-3.99)	1.02	4	1,5
Primary Care Physician	78% (42) (0.65-0.87)	54	4.09 (3.83-4.36)	0.98	4	1,5
Medical/Surgical Specialties	53% (9) (0.31-74)	17	3.53 (3.01-4.05)	1.01	4	2,5
Non-Physician Health Professional	71% (64) (0.61-0.80)	90	3.97 (3.76-4.17)	0.97	4	1,5
Public Health Sciences	70% (80) (0.61-0.78)	114	3.8 (3.58-4.02)	1.19	4	1,5
Nonmedical Reviewer <sup>c</sup>	70% (46) (0.58-0.80)	66	3.94 (3.65-4.23)	1.16	4	1,5

<sup>a</sup>Modified Wald method.

<sup>b</sup>Percentage of family medicine and nonmedical reviewers not significantly different by  $\chi^2$  statistic ( $P = .65$ ).

<sup>c</sup>Means of family medicine and non-medical reviewers not significantly different by 2-tailed unpaired  $t$  test ( $P = .81$ ).

**Conclusions** Reviewers from a wide variety of fields return high-quality reviews, and these reviews are not different from those by experts in the core professional group. Diversity of reviewers can add special perspectives to a generalist medical journal without compromising quality of reviews.

<sup>1</sup>Department of Family Medicine, Box 356390, University of Washington, Seattle, WA 98195, USA, e-mail: wphllps@u.washington.edu; <sup>2</sup>Department of Family Medicine, Case Western Reserve University, Cleveland, OH, USA

**Implementation of a Medical Journal Peer Reviewer Stratification System Based on Quality and Reliability**

Steven Green<sup>1</sup> and Michael Callaham<sup>2</sup>

**Objective** Prior to starting this study 6 years ago *Annals of Emergency Medicine* had a large reviewer pool (N = 989) who demonstrated substantial variability in quality and reliability. We hypothesized that a tiered, dynamic reviewer stratification system might enable our journal editors to target the bulk of their review invitations to our better reviewers and thus improve our efficiency.

**Design** In 2003 we instituted a 3-tiered hierarchical classification and stratified our peer reviewers based on predefined criteria for reviewer quality (ie, average review score) and reliability (eg, response to review invitations, on-time reviews). Using our manuscript management software our editors could then target the bulk of their review invitations to the top performance tier, which constitutes approximately one-fourth of the total. Every 6 months since initiation a senior editor analyzes reviewer performance statistics and promotes or demotes individuals within this dynamic classification. Before-and-after measures of global peer-review efficiency were then assessed.

**Results** We compare 2008 data with 2002, the year prior to the system, and found more top-tier reviewer invitations leading to an on-time review (51% vs 37%), shorter median review turnaround (median, 10 vs 12 days), less late reviews (15% vs 32%), and less reviewers not used in a given year (26% vs 59%). Editors have found the system to be simple and easy to use. No serious problems have been identified. We cannot ascertain how much of the observed improvements are due to this reviewer classification system vs other concurrent quality improvement initiatives including the adoption of our electronic manuscript system a year prior to starting this system.

**Conclusion** Implementation of a tiered, dynamic system stratifying journal peer reviewers by quality and reliability was readily accomplished by *Annals of Emergency Medicine* and has appeared to improve the efficiency of our peer review.

Loma Linda University, 2160 Veloz Dr, Santa Barbara, CA 93108, USA, e-mail: steve@viridissimo.com; University of California San Francisco, San Francisco, CA, USA

**Assessment of Reviewers Recommended by Authors vs Editors: Is There Bias?**

Monica Helton and William Balistreri

**Objectives** To test the hypothesis that author-suggested reviewers (ASRs) are more likely than editor-suggested reviewers (ESRs) to (1) accept invitations to review, (2) recommend acceptance of the manuscript on first review, and (3) concur less frequently with the editor's final decision to accept.

**Design** We retrospectively evaluated the first 300 manuscripts submitted to the *Journal of Pediatrics* that were assigned consecutive manuscript numbers in 2007; 122 manuscripts did not undergo

peer review. For the 188 reviewed manuscripts, we recorded the following: (1) whether the reviewer was suggested by the author or chosen by the editor, (2) the number of ASRs and ESRs who completed reviews, (3) the initial recommendation of the reviewer, and (4) the final decision of the editor. The statistical methods used were the  $\chi^2$  and the McNemar test for correlated proportions. Reviewer recommendations for “acceptance” included “accept” and “accept with revisions.”

**Results** Of the reviewers (n = 873) examined, 37.2% of ASRs accepted the invitation to review (167/449) compared with 41.8% of ESRs (177/424) ( $P = .17$ ). When evaluating reviews, 65.3% of ASRs recommended acceptance (109/167), whereas 54.2% of ESRs recommended acceptance (96/177) ( $P = .04$ ). Editors agreed with 49.5% (54/109) of the accept recommendations of ASRs ( $P < .0001$ ) and with 55.2% (53/96) of ESRs ( $P < .0001$ ).

**Conclusions** There is no evidence to suggest that ASRs are more likely than ESRs to accept an invitation to review. However, ASRs are more likely to recommend acceptance of a submitted manuscript. ASRs and ESRs are more likely to recommend acceptance of a manuscript than an editor. Although this could be due to a variety of factors, including a recommendation of acceptance being paired with one or more recommendations of rejection and priority for publication, this emphasizes the peer review motto of “reviewers advise; editors decide.”

*Journal of Pediatrics*, Cincinnati Children’s Hospital Medical Center, 3333 Burnet Ave, MLC 3021, Cincinnati, OH 45229, USA, e-mail: monica.helton@cchmc.org

**Design** In January 2009 we e-mailed 1 editorial representative from each of 204 journals inviting them to take part in an anonymous online survey. This consisted of 27 multiple-choice questions supplemented by 5 text boxes. The opening question asked respondents to identify which 1 of 6 broad disciplinary categories best described their journals. Subsequent questions addressed various aspects of the peer review process. The survey was intended as a means of quickly gathering approximate answers, rather than necessitating any in-depth research by the respondents.

**Results** A total of 145 journals (71%) were represented. Some of the more interesting findings are presented in **TABLE 7**. All journals employed single-blind or double-blind peer review (some used both). Science and medicine journals showed a markedly higher tendency to use single blinding, to use author suggestions to find reviewers, to rate their reviewers, to share reviewer reports among reviewers, and to notify reviewers of the editor’s final decision than did social sciences and humanities journals.

**Conclusions** These results indicate that a divide between the 2 cultures of humanities and sciences persists and is manifest in patterns of peer review practices. In most, though not all, medical journals are most closely aligned with the sciences, and social science journals align with humanities.

University of Oxford, Oxford University Press, Great Clarendon St, Oxford OX2 6DP, UK, e-mail: huw.price@oup.com

**Peer Review in Journals Published by Oxford Journals**

Huw Price

**Objective** Peer review processes vary widely, and the procedures followed by any given journal are often known only to the editorial team directly involved. We surveyed our journals to get a clearer picture of the range and patterns of peer review practices, experiences, and expectations among the journals that we publish and to facilitate further discussion of best practice with, and between, our journal editors.

**Quality of Peer Reviews in 3 Nursing Journals From the Perspective of Authors and Editors**

Sandra P. Thomas,<sup>1</sup> Mona Shattell,<sup>2</sup> Peggy Chinn,<sup>3</sup> and W. Richard Cowling<sup>2</sup>

**Objective** The literature on review quality in the nursing discipline is quite small. The purpose of this study was to examine the quality of peer review from the perspective of authors and editors in 3 scholarly nursing journals: *Advances in Nursing Science*, *Issues in Mental Health Nursing*, and *Journal of Holistic Nursing*. All 3 journals use double-blind peer review and are indexed in PubMed and CINAHL. Quality of peer review, for the purposes of this study, was defined as constructive guidance for authors to further develop their work for publication and for editors to make sound decisions regarding manuscript disposition.

**Table 7. Peer Review Process Survey Results**

Category	Return rate			Peer review system (n = 145)			Time allocated for review (cumulative) (n = 145)			Methods of finding reviewers (n = 145)		Journals that reward or acknowledge reviewers (n = 132)
	Journals contacted*	Surveys completed	Return rate*	Double blind	Single blind	Mixed	2 weeks or less	4 weeks or less	8 weeks or less	Journal database	Author suggestions	
Overall	204	145	71%	46%	49%	5%	19%	67%	91%	76%	37%	48%
Humanities	54	42	78%	69%	21%	0%	5%	43%	90%	60%	14%	32%
Law	23	12	52%	50%	42%	8%	25%	92%	100%	67%	8%	20%
Mathematics	12	6	50%	0%	100%	0%	0%	17%	33%	33%	17%	0%
Medicine	44	25	57%	28%	72%	0%	48%	92%	96%	92%	68%	83%
Science	37	30	81%	7%	93%	0%	37%	97%	100%	83%	83%	41%
Social science	34	30	88%	77%	17%	7%	0%	50%	87%	90%	13%	67%

\*The number of journals contacted by discipline was estimated based on Oxford Journals’ own classification of all the journals approached. By contrast, returns were assigned to disciplines by the survey respondents themselves.

**Design** A researcher-developed survey instrument was distributed online to all corresponding authors of manuscripts submitted between 2005 and 2007. A total of 319 authors responded (response rate, 69%) under conditions of anonymity. Additionally, one-third of all manuscript reviews completed between 2005 and 2007 (N = 528) were rated by the research team for level of detail, bias, constructive tone, and usefulness to authors in making revisions and for editors in making decisions.

**Results** A majority (73.8%) of authors agreed that reviews by these journals provided constructive guidance, and 75.6% agreed that reviews provided adequate rationale for editors' decisions. Forty percent of authors reported fewer than 10 submissions to any journal, and  $\chi^2$  analysis showed that inexperienced authors perceived review quality less favorably than experienced authors. Critiques of reviews from editorial perspective included insufficient feedback to authors, inconsistency between reviewers' numeric ratings and manuscript disposition recommendations, and occasional evidence of reviewer bias or disrespectful tone.

**Conclusions** The results do not provide compelling evidence to question the worth of the standard peer review approach. Given the relative inexperience of many nurse authors in manuscript submission, it is incumbent upon reviewers and editors to continue providing clear and complete feedback and guidance.

<sup>1</sup>University of Tennessee, College of Nursing, 1200 Volunteer Blvd, Knoxville, TN 37996-4180, USA, e-mail: sthomas@utk.edu; <sup>2</sup>University of North Carolina, Greensboro, NC, USA; <sup>3</sup>University of Connecticut School of Nursing, Storrs, CT, USA

**Nursing Journal Peer Reviewers' Views on Quality Indicators in Publishing**

Molly Dougherty,<sup>1</sup> Margaret Freda,<sup>2</sup> Margaret Kearney,<sup>3</sup> Judith Baggs,<sup>4</sup> and Marion Broome<sup>5</sup>

**Objective** To analyze how nursing journal peer reviewers (NJPRs) identify contributions to nursing in manuscripts, define their priorities in writing reviews, and use journal impact factors (IFs).

**Design** A 69-item online survey was completed in 2007 by NJPRs who were invited by 52 editors of nursing journals worldwide. Editors notified their reviewers of the opportunity to participate once. Anonymous responses were unlinked from journal identification and duplicates removed. Descriptive statistics and  $\chi^2$  were used to test hypotheses that NJPRs familiar with IFs of nursing journals would differ from those not familiar by nursing credential, US vs other country residence, and variables related to research and clinical involvement.

**Results** The NJPRs (N = 1675) were from 44 countries with 74% from the United States, and 90% were nurses. The response rate was 44%. They used contribution to knowledge or research evidence (n = 1404, 83.8%), topic of current interest (n = 1153, 68.8%), and newly emerging area (n = 1141, 68.1%) as indicators of a manuscript's contribution to nursing. In writing their reviews, research rigor (n = 889, 53.1%) and clinical relevance (n = 785, 46.9%) were high priorities. Impact factor was familiar to 810 (48.4%), and 467 (27.9%) used IF to choose journals for submis-

sion of manuscripts. Those familiar with IFs were significantly more often not nurses, not US residents, involved in research, and reviewed most often for a research journal (TABLE 8).

**Table 8. Sample Characteristics and Familiarity With Impact Factor  $\chi^2$  Results (N = 1675)**

Category	n <sup>a</sup>	Response Categories	Familiar With Impact Factor		
			Yes	No or Not sure	
Nurse	1517	Yes	730	669	11.86
		Not Yes	78	40	
US resident	1509	Yes	515	622	95.77
		Not Yes	285	87	
Involved in research	1516	Yes	737	415	217.13
		Not Yes	72	292	
Most often review for clinical journal	1507	Yes	221	422	163.10
		Not Yes	579	285	

<sup>a</sup>Row and column total for  $\chi^2$  analyses; nonresponses not included. Note: df = 1; all  $\chi^2$ , P < .01

**Conclusions** When judging a paper's contribution, NJPRs weigh research and clinical interests. Authors in non-US countries receive monetary and other rewards for articles in high-IF journals, but most NJPRs do not use IF and place clinical considerations ahead of IF. Reviewers who are nurses, US residents, review for clinical journals, or are not involved in research are less familiar with IF.

<sup>1</sup>University of North Carolina at Chapel Hill, School of Nursing, 3602 Andante Dr, Hillsborough, NC 27278, USA, e-mail: m\_dougherty@unc.edu; <sup>2</sup>Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY, USA; <sup>3</sup>University of Rochester School of Nursing, Rochester, NY, USA; <sup>4</sup>Oregon Health Sciences University School of Nursing, Portland, OR, USA; <sup>5</sup>Indiana University School of Nursing, Indianapolis, IN, USA

**Peer Review in Small vs Non-small Biomedical Journals**

Farrokh Habibzadeh<sup>1</sup> and Rob Siebers<sup>2</sup>

**Objective** Although some researchers reported some aspects of the peer review system in some "small journals" (SJs), there is no universally accepted definition for SJ. We hypothesized that the peer review process may be so different in SJs from non-SJs (NSJs) that it can provide a mean to differentiate between SJ and NSJ. We therefore conducted this research to study the peer review process in SJs compared to NSJs.

**Design** On February 25, 2008, members of World Association of Medical Editors (WAME), Eastern Mediterranean Association of Medical Editors (EMAME), Forum for African Medical Editors (FAME), Council of Science Editors (CSE), European Association of Science Editors (EASE), Asociación Mexicana de Editores de Revistas Biomédicas (AMERBAC), and Asociación de Editores de Revistas Biomédicas Venezolanas (ASEREME) were asked to participate in a cross-sectional questionnaire-based online survey. Based on the size of each of these associations and taking into account the common members, we estimate that almost 1000 editors received the invitation. The editors were asked if they thought

their journal is an SJ or not and to describe the review system for their journal. The survey ended 4 months later. Categorical variables were examined by  $\chi^2$ , and the median number of reviewers was compared by Mann-Whitney *U* test.

**Results** During the period 115 editors (62 from SJs and 53 from NSJs) from 30 countries representing 56 journals (42 SJs and 14 NSJs) completed the questionnaire. No duplicate responses were found. The average time from submission to decision stated by both groups was 8 to 12 weeks. Both groups believed that the study design and the methodology section are the most important causes of manuscript rejection. Fifty-four (61%) journals used double-blind systems, 31 (35%) single-blind systems, and 3 open review systems. There was no significant difference between SJs and NSJs.

**Conclusions** SJs are similar to NSJs in terms of most of the peer-review system parameters. However, SJ editors are less likely to use statistical advisors and are more likely to request authors to disclose their conflicts of interest than NSJ editors.

<sup>1</sup>*International Journal of Occupational and Environmental Medicine*, PO Box: 71955-575, Shiraz, 71955 Iran, e-mail: Farrokh.Habibzadeh@gmail.com; <sup>2</sup>Wellington School of Medicine & Health Sciences, Wellington, New Zealand

**Effects of Training Reviewers on Quality of Peer Review: A Before-and-After Study**

Ying YANG

**Objective** To examine effects of short-term training for the reviewers at a Chinese specialized journal on the quality of peer review.

**Design** A before-and-after study included 45 reviewers at the *Chinese Journal of Tuberculosis and Respiratory Diseases* between September 2007 and September 2008. All the reviewers attended a face-to-face training session on peer review for 1 day in February 2008. Three teachers from People's Hospital of Peking University taught the training course. One teacher was a statistician, and the other two teachers were specialists on tuberculosis and respiratory diseases. The training course focused on how to critically appraise research articles and what editors want from reviewers. Three review comments for each reviewer before and after the training were selected randomly and compared. The reviews were evaluated by 1 editorial staff member. The quality of review comments were evaluated by the following: if the comments were returned on time (4 weeks), were 300 words or more, or indicated 4 or more suggestions such as specific errors, detailed suggestions for improvement, or better references. A standardized review draft list was used in this study.

**Results** The mean and range of years of experience of the 45 reviewers are 2.8 (1.5-5.1) years. After the short-term training, the time taken to complete a review increased nearly 20%. Instances of review comments being more than 300 words increased by about 40%. Almost all reviewers could identify the specific errors in the manuscripts. More than half of the reviewers could give suggestions to correct the errors. There was also an increase in the number of reviewers who gave better references (TABLE 9).

**Table 9. Differences Before and After Training on Quality of Peer Review**

Content	Before the Training		After the Training	
	Reviewers, No. (%)	Reviews, No. (%)	Reviewers, No. (%)	Reviews, No. (%)
Time taken to review (<4 weeks)	30 (66)	86 (64)	38 (84)	112 (83)
≥300 words or more than 4 suggestions for improvement	27 (60)	70 (52)	40 (89)	115 (85)
Specific errors identified	25 (55)	65 (48)	41 (91)	118 (87)
Suggestions to correct errors	17 (37)	43 (32)	32 (71)	88 (65)
Better references	4 (9)	10 (8)	15 (33)	39 (31)

**Conclusions** Brief training in peer review appears to improve timeliness and quality of review. The long-term effects still need observation.

Publishing House of Chinese Medical Association, Room 424, 42 Donsi Xidajie, Beijing, 100710 China, e-mail: yangying@cma.org.cn

**Peer Reviewers' Recommendations at the *Journal of General Internal Medicine*, 2004-2008: Style or Substance?**

Richard Kravitz,<sup>1</sup> Peter Franks,<sup>2</sup> Martha Gerrity,<sup>3</sup> Mitchell Feldman,<sup>4</sup> Cindy Byrne,<sup>5</sup> and William Tierney<sup>5,6</sup>

**Objective** To examine factors associated with peer reviewers' initial summary recommendations to reject vs accept or revise submissions to the *Journal of General Internal Medicine*.

**Methods** We analyzed 5881 reviews of 2664 manuscripts submitted between 2004 and 2008. These reviews were performed by 2916 reviewers. The dichotomous dependent variable was the reviewer's initial summary recommendation: reject vs accept/revise. Independent variables included review year, number of reviewers per manuscript, total reviews by each reviewer, time taken for review, article type (eg, original article, perspective, clinical vignette), and a deputy editor's rating of review quality (1 [poor] to 6 [excellent]). We used random effects logistic regression analyses to account for nesting of reviews by manuscript and by reviewer and to calculate the intraclass correlation coefficients (ICC) for manuscripts and reviewers.

**Results** Among 2664 manuscripts sent for peer review, reviewers recommended rejection in 28%. The manuscript-level ICC (between reviewers of each manuscript) was 0.18 (95% confidence interval [CI], 0.13-0.23). With 3 reviews per manuscript, this ICC is equivalent to an alpha reliability coefficient of 0.40 (95% CI, 0.31-47). To achieve an alpha of 0.70, 11 (95% CI, 8-16) reviews would be required. The reviewer-level ICC (across manuscripts) was 0.25 (95% CI, 0.20-0.32). Reviewers more often recommended rejection in earlier years (adjusted odds ratio [AOR], 0.91/year; 95% CI, 0.86-0.97) and when submitting more highly rated reviews (AOR, 1.36/quality unit; 95% CI, 1.28-1.44). Other variables were not significant.

**Conclusions** Between-reviewer agreement on manuscript disposition was modest. Reviewers exhibited greater consistency across

manuscripts, suggesting a stable reviewer style (propensity to reject). While the analysis did not include manuscripts rejected by deputy editors without review nor address the consistency of the narrative portion of the reviews, adequate reliability in summary recommendations will require more reviewers per manuscript and/or more standardization in criteria for summary recommendations.

<sup>1</sup>University of California, Davis, Division of General Internal Medicine, 4150 V St, Suite 2400 PSSB, Sacramento, CA 95817, USA, e-mail: rkravitz@ucdavis.edu; <sup>2</sup>Department of Family and Community Medicine, University of California, Davis, Sacramento, CA, USA; <sup>3</sup>Department of Medicine, Oregon Health Sciences University, Portland, OR, USA; <sup>4</sup>Division of General Internal Medicine, University of California, San Francisco, CA, USA; <sup>5</sup>Regenstrief Institute, Inc, Indianapolis, IN, USA; <sup>6</sup>Division of General Internal Medicine, Indiana University Indianapolis, IN, USA

### Trends in Using Insecure e-Mail Services in Communication With Journal Editors

Krešimir Šolić,<sup>1</sup> Vesna Ilakovac,<sup>1</sup> Ana Marušić,<sup>2</sup> and Matko Marušić<sup>2</sup>

**Objective** Free online e-mail services (eg, Gmail, Yahoo, and Hotmail) are considered to have more security flaws than institutional ones but are widely popular and frequently used. The objective of this study was to analyze the changes in the use of free online e-mail services for correspondence by authors of published papers in a medical journal.

**Design** Contact information of corresponding authors for all papers published in the *Croatian Medical Journal (CMJ)* during a 10-year span (1998-2007) were collected from the *CMJ* electronic archive. Domains of all e-mail addresses were assessed, and contacts were categorized into 4 groups: no e-mail, worldwide available free online e-mail service, free national online e-mail service, and institutional or corporate e-mail address.

**Results** Of 978 authors, 34 had no mail (3.5%), 563 (57.6%) used institutional or corporate e-mail addresses, 246 (25.2%) free national online e-mail service, and 135 (13.8%) worldwide available free online e-mail service. The proportion of authors using world wide available free online e-mail services increased from 7.6% in 1999 to 20.8% in 2007 showing significant increasing trend (Cochrane-Armitage trend test,  $P = .011$ ). Non-Croatian authors ( $n = 520$ , 53%) more often used institutional e-mail addresses than Croatian authors ( $n = 458$ , 47%), ( $\chi^2 = 56.1$ , degrees of freedom = 3,  $P < .001$ ).

**Conclusions** There is a significant increasing trend in using worldwide available free online e-mail services in small general medical journals. Authors should be aware that insecure e-mail services may compromise confidential nature of author-editor communication and should consider institutional e-mail addresses over free ones.

<sup>1</sup>Department of Biophysics, Medical Statistics and Medical Informatics, School of Medicine, Josip Juraj Strossmayer University, J. Huttlera 4, HR-31000, Osijek, Croatia, e-mail: kresimir@mefos.hr; <sup>2</sup>*Croatian Medical Journal* and Split University School of Medicine, Split, Croatia

### Postpublication Perceptions and Influence

#### Canadian Physicians' Perceptions About Biomedical Journals

Erica Frank, Carolina Segura, and Ariella Zbar

**Objective** Those creating major biomedical journals tout their "large audiences" (*Journal of the American Medical Association, JAMA*), claiming to be "truly the 'must read' for all medical professionals" (*New England Journal of Medicine, NEJM*), or "accessed by a vast global audience" (*British Medical Journal, BMJ*). Yet the extent to which these journals are perceived as relevant by rank-and-file physicians is unknown.

**Design** As part of the Canadian Physicians' Health Study (conducted November 2007-May 2008), we queried a random sample ( $N = 3013$ ) of Canadian physicians about their agreement with the following statement: "It is clinically important for me to regularly read the major biomedical research journals." We queried on a 5-point Likert scale from strongly agree to strongly disagree. Comparisons were made in SAS, using  $\chi^2$  tests.

**Results** The majority (66%) of Canadian physicians agree or strongly agree that is clinically important for them to regularly read the major biomedical research journals. This was more true for men than women (68% vs 62%,  $P < .01$ ), non-family physicians than family physicians (74% vs 55%,  $P < .0001$ ), for non-Canadian-born vs Canadian physicians (72% vs 63%,  $P < .0001$ ), for those who attended medical school in another country vs Canada or the US (72% vs 65% vs 61%,  $P = .001$ ), and for those practicing in the inner city vs urban/suburban vs rural/small town/remote settings (75% vs 66% vs 60%,  $P < .0001$ ). Especially high agreement was found among those working primarily in research units (85%) or in academic settings (82%).

**Conclusions** While most academicians and researchers believe that regularly reading major biomedical research journals is important for them, nearly half of Canadian family physicians (who represent about half of Canadian physicians) and rural/small town/remote practitioners do not believe so. Journal editors, publishers, and boards should consider these findings when deciding on content and other publication strategies.

University of British Columbia, School of Population and Public Health, James Mather Building, 5804 Fairview Ave, Vancouver, British Columbia V6T 1Z3, Canada, e-mail: erica.frank@ubc.ca

#### The Convergence of Impact Factor and Media Attention on Discussions Between Doctors and Their Patients: The Case of Rosiglitazone

Jim Nuovo

**Objective** To track the influence of an article published in a high-impact journal on discussions between doctors and their patients in response to concerns about potential medication adverse side effects.

**Design** This was a retrospective analysis of a primary care network's electronic medical record database. From a diabetes registry of 12,246 patients, 369 were identified as taking rosiglitazone prior to the June 14, 2007, publication of an article in the *New England Journal of Medicine*; the article suggesting an increased risk of myocardial events for patients taking the drug. The entire content of all office visits, telephone messages, and medication lists for each patient were reviewed over a 1-year period subsequent to the article's publication. Doctor/patient discussions regarding concerns about rosiglitazone were cataloged including the physician's treatment recommendations.

**Results** There were documented discussions on rosiglitazone's potential adverse effects in 19.2% of this population. All the discussions occurred between June 15 and October 30, 2007. Of this group, 59.2% remained on rosiglitazone. For those advised to continue rosiglitazone, the clinician indicated that he or she wanted more data before determining if the drug was not safe. For those advised to discontinue rosiglitazone, 83.3% were placed on pioglitazone.

**Conclusions** An article suggesting potential adverse effects of rosiglitazone resulted in a documented discussion in 19.2% of patients on this medication. These findings suggest an awareness of this publication by patients, presumably derived from media reports. However, an awareness of this concern did not result in a substantial change in practice. The majority of patients remained on rosiglitazone. The content of these discussions suggest that most physicians recommended waiting for more published data before considering a change. While many factors influence physicians' prescribing behavior, this study demonstrates how an article in a high-impact journal influences the doctor-patient dialogue.

Department of Family & Community Medicine; 4860 Y St, Suite 2300, Sacramento, CA 95817, USA, e-mail: james.nuovo@ucdmc.ucdavis.edu

## Publishing Models

### Transformation of the *American Journal of Obstetrics & Gynecology* From a Traditional Journal to a "New Format"

Thomas J. Garite,<sup>1-3</sup> Roberto Romero,<sup>4,5</sup> Moon Kim,<sup>1,2</sup> Pamela Poppalardo,<sup>6</sup> and the Editors of the *American Journal of Obstetrics & Gynecology*

**Objective** To describe the transformation and results of a major medical journal in response to diminishing readership and major revenue losses that threatened its existence.

**Design** The *American Journal of Obstetrics & Gynecology (AJOG)* is the oldest journal in its specialty with an impact factor ranking it second among general obstetrics/gynecology journals. Subscriptions had dropped from more than 20,000 to 9000 with loss of subscription and advertising revenue. Readership surveys indicated major changes in reading habits of subscribers. In January 2007, *AJOG* transitioned to a new format and from a subscription-only journal to controlled circulation (free) distribution. The print version was changed to include only a 1500-word summary cowritten by a employee scientific writer and the authors, with the full-length article published online.

**Results** In the 2 years following implementation of this change, circulation increased to 44,000 while maintaining two-thirds of its paid subscriptions. Publication time was shortened from 9.9 to 6.9 months. Advertising revenue increased by 45%. According to a PERQ/HIC analysis, there was an increase of high readership from 11% to 29% and of average readership by 18% to 59%. Submissions during this time period increased by 20%. The acceptance rate decreased from 29% to 25% with a small (5%) decrease in articles published. The citation rate increased by 20% and is the highest of any journal in women's health. The impact factor improved from 3.0 to 3.5. The journal's Web site activity increased by 70% in both visits per month and page views. A survey of readers and authors revealed that 85% of readers and 71% of authors found the changes very acceptable or acceptable.

**Conclusion** Major changes in the format and distribution of a 138-year-old journal in response to dramatic changes in reading habits of subscribers resulted in favorable changes in readership, article submissions, and the journal's financial well-being.

<sup>1</sup>Obstetrics and Gynecology, 33025 Ponderosa Trail, Oak Creek, CO 80467, USA, e-mail: tgarite@uci.edu; <sup>2</sup>University of California Irvine College of Medicine; <sup>3</sup>Pediatrics Medical Group, Sunrise, FL, USA; <sup>4</sup>Wayne State University College of Medicine, Detroit, MI, USA; <sup>5</sup>Intramural Division, NICHD, NIH, DHHS, Rockville, MD, USA; <sup>6</sup>Publishing Director, Elsevier, Inc, New York, NY, USA

### A New Journal Mode: Public Peer Review, Open Access on Web 2.0

Xibin SHEN

**Objective** With the improvement of Web service from Web 1.0 to Web 2.0, some of Web 2.0 techniques have enriched information distribution, thereby affecting the academic society. Many academic journals provide new online tools such as interactive Web logs (blogs), bulletin board systems (BBS), and content-sharing sites (Digg) that enable users to be participants rather than information receivers. Users play a principal role in creating, collecting, filtering, organizing, and distributing information materials from different sources.

**Design** Based on this practice, we designed a new mode for creating a journal from the conception of Web 2.0 or a journal 2.0 that is public peer reviewed and open accessed. The journal 2.0 is organized, peer reviewed, revised, and published online by users using Web technology including Web 2.0 or other available techniques. In other words, users play a major role in activities involving science publishing. The users upload, manage, vote, filter, and publish science information on an open platform.

**Results** Journal 2.0 is feasible for filtering low-quality articles by open peer review. Unlike traditional journals, Journal 2.0 uses open peer review by several special designated users in combination with public evaluation. Opening to the public, Journal 2.0 invites several users to make a decision in evaluating manuscripts, which is form-based and involves setting points similar to a gymnastics grading. An original medical article, for instance, can be graded by its relevance and interest (A1), impact (A2), content (A3), originality (A4), and presentation (A5); each item will be multiplied by an assigned weight (W). In addition, users have their own academic

value (V) that is used to assess their academic level or their activities for the journal. In this process, the academic score (S1) from the invited users for each manuscript is a weighted average of the values multiplied by their academic values. So, the academic score of each manuscript during designed peer-reviewing is calculated as the following:  $S_s = ([A1 \times W1 + A2 \times W2 + \dots + A5 \times W5] \times V1 + [A1 \times W1 + A2 \times W2 + \dots + A5 \times W5] \times V2 + \dots + [A1 \times W1 + A2 \times W2 + \dots + A5 \times W5] \times Vn) / n$ . This score plus the user commentaries will be used to primary assess the value of a given manuscript. Meanwhile, another manuscript's evaluation process is open to the public, and their scores (Sp) will mildly adjust the results. In my designed system, however, how the public scores influence the decision can be adjusted by the addition of another weight to special (Ws) or public (Wp) scores. For instance, given an even weight to each section will influence the final value equally to the special ones. At last, the final score of a given manuscript is  $S = S_s \times W_s + S_p \times W_p$ . Within the setting period, the most ranked titles will enter into further steps. The accepted manuscript will return to the author for revising, who can invite his or her colleagues or friends in the platform for editing collaboratively via Wiki. As being done by some traditional journals, published articles will be noted to all users via really simple syndication (RSS) feed, podcast, e-mail alert, etc. In return, the reviewer's academic value will be adjusted by his or her rating for every manuscript and their activities (eg, submission, publishing, providing valuable material of information).

**Conclusion** Under this open circumstance, peer review bias will be eliminated because of its available comment to everyone, and articles will be more acceptable after voting of major users.

*Chinese Journal of Internal Medicine*, 42# Dongsi Xidajie, Beijing, 100710, China, e-mail: cmjshenxb@gmail.com

## Publication Pathway

### What Happens to Rejected Manuscripts?

Kenneth Noller,<sup>1</sup> Sheryl Martin,<sup>2</sup> and James Scott<sup>2</sup>

**Objective** To determine the frequency of later publication of original research papers rejected by the journal *Obstetrics & Gynecology (O&G)*.

**Design** The author search function of PubMed was used to identify papers rejected in 2002 by *O&G* that were later published by any of the authors in another journal. The abstract submitted to *O&G* was compared to that of similar publications, and a numerical value was assigned using 4 variables: authors' names, title, number of study participants, and abstract text. Each variable was coded as exact, close, or dissimilar based on predetermined criteria.

**Results** A total of 194 of 318 (61%) rejected manuscripts were later published. Of these, 73 (38%) were unchanged from the original submission based on exact matches for the 4 variables. Eight (4%) of the 194 were published in journals with an impact factor higher than *O&G*. Sixty-eight (35%) appeared in journals in the top 20 reproductive category based on impact factor. The mean time from *O&G* rejection to publication in another journal was 15.7 months (range, 4-57 months). Thirteen (7%) papers were published in general medical journals, and 83 (43%) in other general

obstetrics-gynecology journals. The remaining articles appeared in other specialty, obstetrics-gynecology subspecialty, epidemiology, nursing, and review journals. Of the 124 papers (39%) that were not published in any journal, 115 were observational studies. There were 5 randomized controlled trials and 4 papers that did not fit the original research category.

**Conclusion** Many papers originally rejected by *Obstetrics & Gynecology* are later published in other peer-reviewed journals. These results have important implications for all scientific journals. The values and quality of the original peer review, whether all rejected articles should be published, impact on scientific validity, and effect on the increasing information overload for clinicians are issues that should be addressed.

<sup>1</sup>Department of Obstetrics and Gynecology, Tufts University School of Medicine, Boston, MA, USA; <sup>2</sup>*Obstetrics & Gynecology*, Department of Obstetrics and Gynecology, University of Utah Medical Center, 423 Wakara Way, Suite 201, Salt Lake City, UT 84108-1242, USA, e-mail: james.scott@hsc.utah.edu

### Submission and Peer Review Frequencies Before Acceptance: An Analysis of Submissions by Mayo Clinic Section of Scientific Publications, 2006

Colleen M. Sauber, LeAnn M. Stee, Sarah M. Jenkins, and Margery J. Lovejoy

**Objective** To estimate the percentage of manuscripts submitted to peer-reviewed journals that undergo peer review at up to 3 journals before publication.

**Design** We hypothesized that 65% of manuscripts submitted to medical and scientific journals undergo peer review 1 or 2 times, and at least 80% undergo review up to 3 times before publication in the identified journal. We retrospectively analyzed manuscripts whose submissions were tracked by Mayo Clinic's publications section in 2006. Data on number of revisions, rejections, and acceptance were recorded for a simple random sample of all manuscripts undergoing full-service process and submitted to peer-reviewed journals that year. Frequency and percentage of revisions, rejections, and acceptances were summarized over all submissions and the random sample. Observed percentages and exact binomial 95% confidence intervals (CIs) were compared with hypothesized values. A journal submission response was considered a peer review by journal and outside experts.

**Result** The sample comprised 104 manuscripts and 217 submissions to 129 journals. Five manuscripts (4.8%; 8 [3.7 %] submissions) lacked data beyond final submission. Of the manuscripts, 73 (70.2%) underwent 1 or 2 submissions, thus, 1 or 2 peer reviews; 86 (82.7%) underwent up to 3 peer reviews (TABLE 10). Observed percentages were consistent with hypothesized values. Among clinical articles (n = 80), 67 manuscripts (83.8%) underwent up to 3 submissions; all 13 case reports had 3 or fewer submissions. Among all manuscripts, 73 (70.2%) were ultimately accepted, and 36 (34.6%) were accepted after submission to only 1 journal. Of these 36 manuscripts, 6 were accepted with no revision. Median time from submission to acceptance was 224 days (range, 3-1571 days).

**Table 10. Total Submissions for Each Unique Manuscript Among 104 Randomly Selected Manuscripts Submitted in 2006 and the Clinical Articles and Case Reports of the Sample**

Total No. of Submissions	All Manuscripts (N = 104)		Clinical Articles (n = 80)		Case Reports (n = 13)	
	No. (Cum %)	95% CI	No. (Cum %)	95% CI	No. (Cum %)	95% CI
1	56 (53.8)	43.8-63.7	42 (52.5)	41.0-63.8	10 (76.9)	46.2-95.0
2	17 (70.2)	60.4-78.8	14 (70.0)	58.7-79.7	2 (92.3)	64.0-99.8
3	13 (82.7)	74.0-89.4	11 (83.8)	73.8-91.1	1 (100)	75.3-100
4	7 (89.4)	81.9-94.6	5 (90.0)	81.2-95.6		
5	8 (97.1)	91.8-99.4	6 (97.5)	91.3-99.7		
6	1 (98.1)	93.2-99.8	1 (98.8)	93.2-100		
7	2 (100)	96.5-100	1 (100)	95.5-100		

Abbreviations: CI, confidence interval; cum %, cumulative percentage.

**Conclusions** During submission to peer-reviewed journals, approximately 80% of manuscripts undergo up to 3 peer reviews. Peer review influences timing of publication of scientific data and discussion. Focusing submission to an appropriate journal could decrease time to publication.

Mayo Clinic, Section of Scientific Publications, 200 First St SW, Rochester, MN 55905, USA, e-mail: sauber.colleen@mayo.edu

### Quality of the Literature

#### The Evolution of Thoracic Surgical Literature

Maurice Blitz,<sup>1</sup> Andrew Graham,<sup>2</sup> Sean P. McFadden,<sup>2</sup> Sean C. Grondin,<sup>2</sup> and Gary Gelfand<sup>2</sup>

**Objective** Modern surgical care is increasingly driven by evidence-based decision making. In an effort to assess the quality of evidence available to thoracic surgeons the levels of evidence of original research and meta-analyses published in the top 3 thoracic surgery subspecialty journals (identified using impact factor) were determined and compared for the years 1998, 2002, and 2006.

**Design** Searches of the table of contents of the *Journal for Thoracic and Cardiovascular Surgery*, the *Annals of Thoracic Surgery*, and the *European Journal of Cardio-Thoracic Surgery* were performed. All original research or meta-analyses pertaining to thoracic surgery for the years 1998, 2002, and 2006 were identified. The individual abstracts and their corresponding manuscripts were then evaluated and classified by 2 independent observers using the levels of evidence from the Oxford Centre for Evidence-Based Medicine. For each publication within each specified year the distribution of the levels of evidence of the manuscripts was compared.

**Results** Six hundred sixty-three manuscripts were reviewed; 39 manuscripts were categorized as level 1 evidence (randomized and controlled trials and systematic reviews of randomized and controlled trials). Four (2.4% of all included publications) were identified in 1998, 12 (4.9%) in 2002, and 21 (8.3%) in 2006. This trend to a higher proportion of the publications determined to be level 1 evidence is significant ( $P = .009$ ). Initial review resulted in a Cohen's kappa of 0.335, suggesting very poor agreement. When the data was categorized as level 1 or levels 2 to 5 (not level 1),

the kappa value of 0.695 (95% confidence interval [CI], 0.507-0.884) and the corresponding 97.2% (standard error +/- 9.6%) agreement was much more acceptable.

**Conclusions** The overall trend toward the publication of manuscripts deemed to be of higher levels of evidence has increased from 1998 to 2006. Unfortunately, the proportion of manuscripts determined to be level 1 evidence is low. More emphasis needs to be placed on the completion of systematic reviews as well as on the design, implementation, and publication of randomized and controlled trials so that thoracic surgical literature can maintain its standing within the evidence-based milieu of the general medical literature.

<sup>1</sup>St Joseph's Health Centre, 30 The Queensway, SSW 245, Toronto, Ontario M6R-1B5, Canada, e-mail: mobilitz@shaw.ca; <sup>2</sup>Division of Thoracic Surgery, Department of Surgery, University of Calgary, Calgary, Alberta, Canada

#### Pharmacovigilance and Local Literature: An Italian Bibliography

Claudio Oliveri and Daniela Ranzani

**Objective** This study explores existing scientific and medical Italian literature in order to classify local publications and to create a database usable in postmarketing pharmacovigilance activities, as prescribed by the European Guidelines on Pharmacovigilance for Medicinal Products for Human Use (Volume 9A).

**Design** Local refers to the type of dissemination of the journals: only nonindexed Italian publications have been considered, assuming a national circulation. The publications were selected in 2 different ways. A keyword-based search was performed in Ulrich's Database. We subsequently examined the official organs of 195 societies, members of the Italian Federation of Medical Societies (Federazione Italiana Società Medico Scientifiche, FISM). This study was conducted in March 2009. The 2 data sets were first analyzed individually and then compared and combined to form a unified database.

**Results** A total of 344 Italian scientific medical journals were identified: 176 (51%) journals were indexed in international databases [40 (23%) had an impact factor], and 168 (49%) were nonindexed journals, which could be defined as local. It was found that among nonindexed journals, 117 (70%) were published in Italian only. The nonindexed journals cover all major therapeutic areas; the most represented areas included neurology/psychiatry (24 journals), surgery (13), cardiovascular medicine (11), dentistry (11), pediatrics (9), and diagnostic (9). There were 21 nonspecialized journals.

**Conclusions** A significant portion of the scientific data published within Europe originates within non-English local journals that may lack an impact factor and not be indexed in international databases, causing this data to be inaccessible at a global level. The creation of a national registry is the first step to allow pharmaceutical companies to screen local bibliographical sources according to rules and regulations governing medicinal products in the European Union (Eudralex). The methodology used to compile our database can be extended to other European countries.

Wolters Kluwer Health Italy Ltd, Via Lanino 5, 20144, Milan, Italy, e-mail: claudio.oliveri@wolterskluwer.com

## Quality of Online Information

### Multilayer Quality Control for Publications and Online Lectures

Faina Linkov,<sup>1</sup> Gilbert S. Omenn,<sup>2</sup> Ismail Serageldin,<sup>3</sup>  
Vinton Cerf,<sup>4</sup> and Ronald LaPorte<sup>1</sup>

**Objective** Quality control (QC) of scientific of educational presentations online is a serious concern for all scientific disciplines. Peer review, the golden standard of quality in various scientific disciplines, may not be optimal for the review of online lectures because it is labor intensive and has low throughput. This paper will discuss the Supercourse, global library of more than 3600 online lectures available at [www.pitt.edu/~super1](http://www.pitt.edu/~super1) and several alternative quality control approaches that are being developed as part of this global effort.

**Design** In the Supercourse, we are utilizing both traditional quality control systems (lecture review forms) and novel approaches by developing or adopting multiple quality assessment methods from other fields for PowerPoint online. These systems include expert editorial board, screening to identify inappropriate lectures, opinion of experts, personal characteristics of authors, Web statistics for utilization of lectures (hits, links, page rankings), keynote speeches (benefiting from the choice of speakers by the relevant society), publications and citations from Google Scholar, and a model similar to the National Institutes of Health style review. The lecture review forms at the end of each module rates the lecture on content, presentation, relevance, and overall rating using a 5-point Likert-like scale, with more than 14000 review forms accumulated in the past 8 years. We calculated correlation coefficient between expert and nonexpert reviewers for the first 10 lectures that received the maximum number of reviews.

**Results** The mean overall score for all lectures was +/- 4.54 out of a possible 5. More than 50% of the Supercourse reviewers were "expert reviewers" in that they were medical doctors and university professors. The correlation coefficient between reviews of experts and nonexperts was 0.79 ( $P < .05$ ) for the first 10 lectures, demonstrating that both experts and nonexperts rate lectures highly. Other quality metrics utilized in the Supercourse also demonstrate that experts rate lectures very highly.

**Conclusions** The multilayer approach to QC of online materials is a novel approach that incorporates existing QC mechanisms, such as peer review, as well as explores the utilization of alternative methodologies. We are currently working on making all of the metrics and findings from the QC evaluation available to the end user and on evaluating the effectiveness of all quality metrics. Our hope is that future scientific research on peer review as well on emerging multilayer QC methodologies will help us to determine best measures of QC.

<sup>1</sup>University of Pittsburgh Cancer Institute, Division of Cancer Prevention and Population Science, 5150 Centre Ave, Suite 4-C, Room 466, Pittsburgh, PA 15232, USA, e-mail: [faina.linkov@gmail.com](mailto:faina.linkov@gmail.com); <sup>2</sup>University of Michigan, Ann Arbor, MI, USA; <sup>3</sup>Library of Alexandria, Alexandria, Egypt; <sup>4</sup>Google, Mountain View, CA, USA

### Citations to Web Sites in Scientific Papers: The Permanence of Native and Archived References

Andrea Thorp<sup>1</sup> and David Schriger<sup>2</sup>

**Objective** Uniform resource locators (URLs) are being cited within the medical literature with increasing frequency. In contrast to cited printed material, a cited Web page is not permanent. Literature suggests that half of the Internet references are inaccessible after 1 year. One solution to ensure access to the intended information is to archive the Internet reference prior to publication. WebCite is a free archiving service that creates a "snapshot" of the desired Web page and provides a permalink URL for that snapshot.

**Design** Internet references included in original research articles available in the online publication of *Annals of Emergency Medicine* from June 2007 to February 2009 were included for analysis. Each original referenced URL was archived using [webcitation.org](http://webcitation.org). To evaluate the permanence of the original and archived Web pages we attempted to access each at 6-month intervals after online publication.

**Results** Data collection is ongoing. Interim analysis includes 98 original/archived WebCite pairs checked a mean of 268 (standard deviation [SD], 192) days (median, 310) days after online publication. Ninety-nine percent (97/98) of the archived WebCite URLs readily accessed the correct Web page. The 1 inaccessible WebCite linked to the correct URL but had no content visible. In contrast, only 62% (61/98) of original URLs were available in original form including 74% checked within 6 months, 56% checked within 6 to 12 months, and 52% checked within 12 to 24 months of publication. Twelve URLs were completely gone (a "404 error"); the other 25 had changed over time. Four of the Webcite URLs had failed to capture the entire page.

**Conclusions** Archiving Internet references may offer a more permanent means to retrieving information cited from a Web page. The scientific community should consider establishing a mechanism for archiving Internet references in medical publications.

<sup>1</sup>Loma Linda University Medical Center, Pediatric Emergency Medicine, 11234 Anderson St, MC-A108, Loma Linda, CA 92373, USA, e-mail: [athorp@llu.edu](mailto:athorp@llu.edu); <sup>2</sup>UCLA Medical Center, Emergency Medicine Center, Los Angeles, CA, USA

### Understanding Why the US National Library of Medicine Fails to Properly Index the Publication Type of a Number of Randomized Controlled Trials

Susan Wieland<sup>1</sup> and Kay Dickersin<sup>2</sup>

**Objective** The indexing of randomized controlled trials (RCTs) in databases (ie, as RCT[pt] in MEDLINE) facilitates efficient searching for trials. We examined 591 RCTs added to PubMed in 2005, but not tagged RCT[pt] in MEDLINE, to understand potential explanations for nonindexing.

**Design** Study reports were initially identified through the Cochrane MEDLINE retagging project, ending in 2006. The project used a sensitive search strategy to search MEDLINE for RCT reports not tagged as RCT[pt] and hand-searching to identify RCTs. In this analysis, 2 independent reviewers read the title and abstract of each indexed RCT not tagged RCT[pt] and classified journal characteristics, type of report (eg, secondary analysis, design, and rationale), and National Library of Medicine topic indexing. Reports could be classified as more than 1 type. Both reviewers completed preliminary classification of the first 100 records, with perfect agreement on RCT status and collaborative development of a classification scheme. One reviewer completed classification of all 591 records, and we will report on reviewer classifications and agreement for all 591 at the Congress.

**Results** Twenty-two percent (21/97) of confirmed RCTs appeared to be straightforward RCTs. The most common report types otherwise were secondary analyses of RCT data (31/97), the rationale, design, pilot, or baseline data for an RCT (21/97), and observational analyses from RCT data (14/97). Four percent (4/97) were tagged Clinical Trial[pt]. Half the records (45/97) were MeSH tagged "Randomized Controlled Trials," indicating that RCTs were the topic.

**Conclusions** Preliminary analysis indicates that nearly all RCT records not indexed RCT[pt] in MEDLINE and identified by Cochrane do describe trials. Authors and editors should ensure that the term "randomized controlled trial" is used in their title or abstract, regardless of presence of or type of analysis. Searchers should use text words for randomization and the MeSH "Randomized Controlled Trials" in addition to RCT[pt] when searching MEDLINE, especially for records added after 2005.

<sup>1</sup>Brown University, Box G, Providence, RI 02912, USA, e-mail: lswieland@yahoo.com; <sup>2</sup>Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology, Baltimore, MD, USA

## Quality of Reporting

### Quality Review of Clinical Guidelines

Anne Hilde Rosvik, Trond Bjornerud, Espen Movik, and Magne Nylenna

**Objective** To test the feasibility and the benefit of quality review of clinical guidelines and to compare this process with traditional peer review of scientific articles.

**Design** Of 352 clinical guidelines included in the Norwegian Electronic Health Library, 112 were suitable for assessment with Appraisal of Guidelines Research and Evaluation (AGREE), an evaluation instrument. Inclusion criteria for assessment were Norwegian origin of guideline, published in the period 2000-2008, comprehensive guidelines (not procedures), national coverage, and direct relation to patient care. AGREE contains 23 items organized in six domains, each assessed on a scale from 1 (strongly disagree) to 4 (strongly agree). Domain scores were calculated by summing up all the scores of the individual items in a domain and by standardizing the total as a percentage of the maximum possible score for that domain. Eighteen doctors and nurses

were trained and certified as reviewers, and each guideline was assessed by 2 reviewers. Assessments were carried out during 2007 and 2008.

**Results** Each reviewer used from 2 to 5 hours on each guideline. The average cost for evaluating 1 guideline amounted to 6400 NOK (800 Euro) including administration. The average scores ( $\pm$  standard deviation) for the 6 domains are as follows: scope/purpose: 72% ( $\pm$  16); clarity/presentation: 69% ( $\pm$  17); stakeholder involvement: 46% ( $\pm$  19); editorial independence: 36% ( $\pm$  19); applicability: 34% ( $\pm$  16); and rigor of development: 33% ( $\pm$  23). By combining all domains 12 guidelines received the conclusion "strongly agree" (most scores above 60%), 91 "recommend with provisos or alterations" (most scores between 30% and 60%), and 8 "would not recommend," and 1 "unsure."

**Conclusions** Most Norwegian guidelines do not fulfill the quality criteria. The final conclusion "recommend with provisos or alterations" given to 4 of 5 guidelines is not very helpful. To get useful information about the guideline quality, clinicians have to look at each domain score. Publication of AGREE results, together with each guideline, have, however, made the guideline quality more transparent and easily accessible for clinicians and will hopefully increase quality over time.

Norwegian Knowledge Centre for the Health Services, The Norwegian Electronic Health Library, PO Box 7004, St Olavs plass, Oslo 0130, Norway, e-mail: aro@helsebiblioteket.no

### Development of Clinical Guidelines in Norway: Do Patients Participate?

Anne Hilde Rosvik, Trond Bjornerud, Espen Movik, and Magne Nylenna

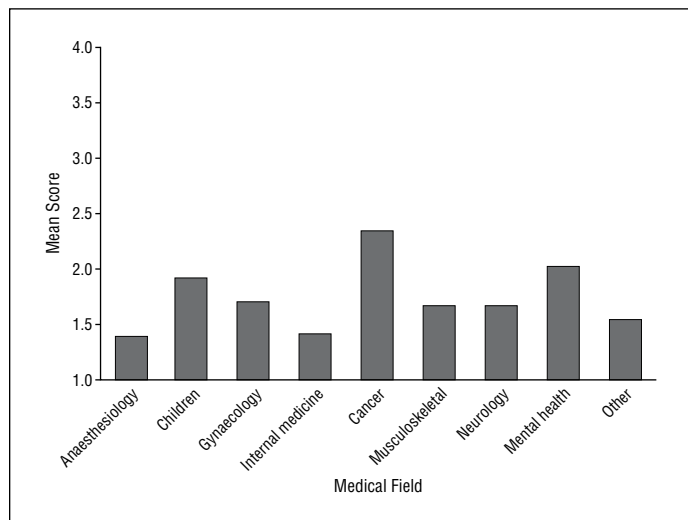
**Objective** It is widely accepted that patients' views and perspectives should be an integrated part of clinical guidelines. We aimed at exploring to what extent patients participate in guideline development in Norway.

**Design** The Norwegian Electronic Health Library includes 352 clinical guidelines freely available online, currently used in clinical practice, in the Norwegian language, published in the period 2000-2009, and developed by accepted and well-known organizations or institutions. A total of 112 guidelines were assessed with Appraisal of Guidelines Research and Evaluation (AGREE) during 2007 and 2008, an instrument for evaluating guidelines. Inclusion criteria for assessment were Norwegian origin of guideline, comprehensive guidelines (not procedures), national coverage, and direct relation to patient care. AGREE covers 23 different items organized in 6 domains. We analyzed the item "The patients' views and preferences have been sought" in the domain "Involvement of stakeholders," which was assessed by 2 appraisers, each rating the issue on a scale from 1 (strongly disagree) to 4 (strongly agree).

**Results** The patient-participation item had a mean score of 1.8 (confidence interval [CI], 1.6-2.0). In 47 (41%) of the guidelines there had been no patient involvement at all. The patient involvement item scored significantly lower than involvement of stakeholders in general. Guidelines produced by the health authorities

scored higher than guidelines produced by professional organizations, mean 2.3 (2.0-2.6) vs 1.5 (1.2-1.8). Guidelines on cancer, mental health, and children had the highest scores (FIGURE 4). There was no significant difference between older (2000-2005) and newer (2006-2009) guidelines. Twelve guidelines scored a top 4 from at least 1 of the appraisers. Of these 10 guideline drafts had been sent to patient organizations for feedback and comments, and in 5 cases patients or patient representatives had participated in the development group.

**Figure 4. Mean Scores for Patient Participation in Guideline Development According to Medical Field of Guideline**



**Conclusions** Patients are generally not participating in development of guidelines despite the intention. There is no improvement over the last years.

Norwegian Knowledge Centre for the Health Services, The Norwegian Electronic Health Library, PO Box 7004, St Olavs plass, OSLO 0130, Norway, e-mail: aro@helsebiblioteket.no

**Variability of Standards for Reporting Tumor-Graft Data in Preclinical Cancer Therapeutic Studies**

Adam Pascoe,<sup>1</sup> Elizabeth Sugar,<sup>2</sup> Scott Kern,<sup>1</sup> and Nilofer Azad<sup>1</sup>

**Objective** To characterize the methodological and statistical parameters for the reporting of tumor-graft experiments for oncologic drug development.

**Design** Using 2007 impact factors, we identified the most-cited medical (3) and oncology journals (5) with tumor-graft reports. All publications used tumor-graft implantation in murine models. For each experiment, the characteristics of the animal models, tumor grafts, experimental therapy, and statistical analysis were examined.

**Results** We examined 145 articles describing 255 experiments from October to December 2008. The papers spanned a wide range of disease types, graft locations, and treatments. Missing data for key design variables was found in 51% of experiments. There was no

standard for the treatment initiation of each experiment, commencing based on elapsed time from tumor implantation (63%) or on tumor volume (23%). Eleven percent of the papers did not report a negative control. There was significant variation among outcome measurements: 70% used tumor volume; others used animal survival (20%), postsacrificial examination (21%), and/or biological changes in the tumor (29%). Tumor volume was evaluated via volume, diameter, area, imaging, and palpation without a standard. Volume was derived using 11 specified formulae or an unspecified calculation (22%). Statistical tools used included the *t* test (47%), ANOVA (21%), and log-rank test (18%); however, 23% of the studies did not report a statistical method for evaluating data. *P* values were reported in 70% of the experiments and in 8% of the manuscripts' abstracts. Ninety-four percent of studies were reported as positive.

**Conclusions** Tumor-graft studies are reported without a standard, often without the methodological information necessary to repeat and confirm the experiments. The high percentage of positive trials indicates potential publication bias. Considering the widespread use of such experiments in choosing drugs for oncology clinical development, we feel it is important for scientists and publishers to develop a consensus set of publication guidelines for reporting experimental and statistical procedures, and we present our initial recommendations.

<sup>1</sup>Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, 1650 Orleans St, Room 4M10, Baltimore, MD 21231, USA, e-mail: nazad2@jhmi.edu; <sup>2</sup>Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

**Challenges in Peer Reviewing Evidence-Based Clinical Practice Guidelines: Do We Know the Degree to Which Guideline Recommendations Reflect Underlying Evidence in Ophthalmology?**

Tianjing Li, Roberta Scherer, Elizabeth Ssemamanda, Ann Ervin, and Kay Dickersin

**Objective** To compare evaluation of the 2005 American Academy of Ophthalmology (AAO) Glaucoma Preferred Practices Patterns (PPPs) guidelines, using the Appraisal of Guidelines Research & Evaluation (AGREE) instrument, with the actual evidence, to understand the challenges of guideline peer review.

**Design** Two people independently used AGREE to assess the rigor of development of the AAO PPPs. Since AGREE does not provide a mechanism for assessing whether the evidence cited in a guideline is appropriate, we identified evidence that would have been available and assessed the methodological quality of relevant systematic reviews. In this context, we searched the Cochrane Library, PubMed, and EMBASE for systematic reviews, and CENTRAL for randomized controlled trials (RCTs). Two people independently reviewed search results, selected studies, and assessed systematic review quality.

**Results** The AGREE score for rigor of development was 48%. Items rated low were searching, criteria for selecting evidence, methods used for formulating the recommendation, and external review. The AAO PPPs cited 5 systematic reviews and 28 RCTs, whereas

we identified 18 systematic reviews and more than 300 potentially eligible RCTs that would have been available as of the last date of guideline literature search (April 2004). Of the 12 reviews in which we assessed quality (reasons for not assessing quality in 6 reviews: multiple publication,  $n = 4$ ; non-English,  $n = 1$ ; full text unavailable,  $n = 1$ ), 5 had predetermined eligibility criteria, 4 searched more than 1 database, and 6 evaluated the quality of included studies. Four reviews fulfilled all methodological criteria.

**Conclusions** Although a global appraisal of the rigor of development of guidelines identifies issues for gathering and synthesizing evidence, it does not inform readers the degree to which these recommendations reflect the underlying evidence. Discussion of the role of peer review in guideline publication is needed, given guidelines' pivotal role in knowledge translation and health care decision making.

Johns Hopkins Bloomberg School of Public Health, Mailroom W5010, 615 N Wolfe St, Baltimore, MD 21205, USA, e-mail: ti@jhsph.edu

#### Reporting Guidelines for Surveys: Limited Guidance and Little Adherence

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

**Objectives** To identify articles providing guidance on reporting mail surveys (self-administered questionnaire data) in the health science literature and to compare recommended practice with reported practice.

**Design** A search strategy was conducted in MEDLINE and PsycINFO electronic databases. Further strategies to identify relevant papers included reviewing reference lists of included studies, using the related articles feature in PubMed for all papers meeting our eligibility criteria, and reviewing relevant textbooks and Web sites. Eligible papers were those written in English and (1) provided guidance on the reporting of survey research or (2) reported evidence of the quality of reporting of survey research. For each source providing guidance on the reporting of survey research, the number of items included and consensus across the guidelines were evaluated. For papers presenting evidence of reporting practice, each aspect of survey reporting addressed was extracted. The data were summarized descriptively.

**Results** Four papers and 1 Web site ([www.aapor.org](http://www.aapor.org)) that provided guidance on the reporting of survey research in the form of a numbered list or checklist were identified; however, none were validated instruments. One checklist was specific to the reporting of internet surveys. For the remaining sources, 39 different reporting items were identified, but only 3 items (description of the research tool, representativeness of the sample population, and response rates) appeared in all guidelines. Six papers were identified that assessed the quality of reporting of some aspect of survey research. Three papers looked at response rate reporting, 1 study evaluated the reporting of nonresponse analysis, and 2 papers assessed description of or access to the questionnaires used in survey research. Overall, these papers indicated that there was suboptimal reporting of these domains.

**Conclusions** There is limited guidance and no consensus regarding the optimal reporting of survey research. Those recommended reporting criteria that have been evaluated are poorly reported.

Ottawa Health Research Institute, Clinical Epidemiology Program, ASB 2-013, 1053 Carling Ave, Ottawa, Ontario K1Y 4E9, Canada, e-mail: cbennett@ohri.ca

#### Publication of Methods in Focus Group Research: Are There Quality Standards for Qualitative Research?

Chima D. Ndumele,<sup>1</sup> Genna Ableman,<sup>2</sup> Beverly Russell,<sup>3</sup> Edith Gurrola,<sup>2</sup> and LeRoi S. Hicks<sup>2</sup>

**Objective** To identify described strategies and success rates in recruiting non-Hispanic black and Latino participants in published focus group research.

**Design** We conducted a systematic review of 264 focus groups in 40 published articles between 1996 and 2009 using the PubMed/MEDLINE and CINAHL databases. Articles were excluded if groups did not target adult minority chronic disease patients, were conducted internationally, or if the primary recruitment was not for focus groups. Patients in studies had a diagnosis of or were at high risk for either coronary disease, diabetes, hypertension, or asthma.

**Results** The majority of studies (36/40) described recruitment strategies in detail; however, fewer studies reported receiving institutional review board approval (65%) or informed consent from participants (63%). The most commonly reported methods of patient recruitment were active in-person (30%), mailed letters (25%), and flyers (25%); 17/40 (43%) studies reported using a combination of methods. The most common recruitment location for studies was health centers (57%) followed by churches (13%). The majority of published studies (80%) did not fully report how successful recruitment efforts were. Only 8 studies reported the total number of potentially eligible persons initially contacted and described the proportion who agreed to participate in the study. Among the remaining studies, 11 (34%) reported the proportion of patients who attended the focus groups compared to those who agreed to participate; these percentages ranged from 32% to 85%. Overall, 50% (6/12) of studies with  $\geq 60\%$  participant retention reported offering participants monetary incentives, while only 33% (4/12) of studies reported using follow-up phone calls between the participants initial agreement and the focus group.

**Conclusions** There remains no consensus regarding the relative effectiveness of strategies due to a lack of reporting on the results of study recruitment efforts. Journal editors and reviewers should consider promoting a universal standard for reporting methodology in focus group manuscripts.

<sup>1</sup>Alpert School of Medicine, Department of Community Health, Brown University, 121 S Main St, 2nd Floor, Providence, RI 02912, USA, e-mail: chima\_ndumele@brown.edu; <sup>2</sup>Harvard Medical School, Boston, MA, USA

### Inconsistent and Incompletely Reported Outcomes in Randomized Trials of Kidney Transplantation

Angela C. Webster,<sup>1,3</sup> Loren P. Ruster,<sup>1</sup> Alistair Merrifield,<sup>1</sup> Gail Y. Higgins,<sup>2</sup> Daniel S. Owers,<sup>1</sup> and Jonathan C. Craig<sup>1,2</sup>

**Objective** Comparing treatment interventions is challenging when published trial outcomes are not consistent. The optimum result after kidney transplantation is a live patient with a transplant that functions well. We systematically reviewed kidney function measures reported in contemporary randomized trials of kidney transplantation.

**Design** Using the Specialised Register of the Cochrane Renal Group we identified trials of immunosuppression in kidney recipients, 2000-2007. Two authors worked independently using standardized tools, with differences resolved by discussion with a third author. We abstracted details of outcomes death, transplant loss, and measures of relative transplant function (creatinine and estimates by formula, eGFR). We examined completeness of published data and identified factors associated with complete reporting using logistic regression.

**Results** Of 99 trials 94 reported patient death, 91 transplant loss, 74 creatinine, and 56 eGFR. For creatinine, 59 reported mean value at trial endpoint, 16 median, 2 mean-change, and 2 number with value greater than threshold. Fifty reported mean eGFR at trial endpoint, 8 median and 2 mean change calculated using at least 4 different formulae (7 method not stated). For creatinine, 52 reported incomplete data, and eGFR 33, most commonly missing number participants contributing measurements (dependent on continued transplant function and available creatinine value at nominated time point) and/or estimate precision. The odds of complete reporting of data were significantly increased for trials sponsored by the pharmaceutical industry for death ( $P = .0025$ ), transplant loss ( $P = .03$ ), and creatinine ( $P = .02$ ), but journal audience, journal impact factor, year of publication, geographical location of trial, and number of participants did not significantly influence data completeness.

**Conclusions** Inconsistent definition and incomplete reporting of outcomes in trials impedes meaningful comparison of results, may mislead users of evidence, and promotes selective over systematic use of available evidence. Standardized definitions and enforced reporting standards would counter this.

<sup>1</sup>University of Sydney, School of Public Health, Edward Ford Building A27, Sydney, NSW 2006, Australia, e-mail: angela.webster@gmail.com;

<sup>2</sup>Children's Hospital at Westmead, NSW, Australia; <sup>3</sup>Centre for Renal and Transplant Research, Westmead Hospital, NSW, Australia

### The SPIRIT Initiative: Defining Standard Protocol Items for Randomized Trials

An-Wen Chan,<sup>1</sup> Jennifer Tetzlaff,<sup>2</sup> Douglas G. Altman,<sup>3</sup> Peter Göttsche,<sup>4</sup> Asbjørn Hróbjartsson,<sup>4</sup> Karmela Krleža-Jerić,<sup>5</sup> Andreas Laupacis,<sup>6</sup> and David Moher<sup>7</sup>

**Objective** The study protocol for a randomized trial serves as the origin of subsequent trial conduct and reporting. With recent

international shifts in policy and legislation toward increased public access to information from trial protocols, these documents will become increasingly important for transparency and critical appraisal of trial methods and results. However, the completeness of protocols varies greatly, partly due to variable standards for their content. We aim to develop evidence-based recommendations for core items to address in the protocol of a randomized trial.

**Design** An evidence-based checklist of key items to include in a trial protocol is being developed through 3 processes: (1) systematic review of existing guidelines for protocol content identified from MEDLINE, EMBASE, Cochrane Library, book chapters, and citation snowballing; (2) systematic review of empiric evidence supporting the importance of addressing specific protocol items (searches of MEDLINE, EMBASE, and the Cochrane Library, as well as citation snowballing were used); (3) Delphi consensus process involving 3 rounds of e-mail survey (n = 96 participants from 17 countries) and a consensus meeting (n = 18 participants) (participants were selected using purposive sampling to ensure representation from key stakeholder groups).

**Results** We identified 27 guidelines for protocol content, only 1 of which was specific to randomized trials. None were developed using an evidence-based approach. More than 1500 articles have also been systematically reviewed to identify empiric evidence for specific protocol items. Following the iterative Delphi consensus process, the current draft checklist consists of 35 items in 5 categories.

**Conclusions** The evidence-based SPIRIT recommendations will benefit peer reviewers, researchers, and other key stakeholders by helping to standardize the core content of trial protocols and improve their quality, facilitate the critical appraisal of trials, and ultimately enhance transparency in clinical trials research.

<sup>1</sup>Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA, e-mail: chan.anwen@mayo.edu; <sup>2</sup>Chalmers Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada; <sup>3</sup>Centre for Statistics in Medicine, Wolfson College Annexe, Oxford, UK; <sup>4</sup>Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark; <sup>5</sup>Canadian Institutes of Health Research, Ottawa, Ontario, Canada; <sup>6</sup>St Michael's Hospital, Li Ka Shing Knowledge Institute, Toronto, Ontario, Canada; <sup>7</sup>Clinical Epidemiology Methods Centre, Ottawa Health Research Institute, The Ottawa Hospital, Ottawa, Ontario, Canada

### Assessment of Randomized Controlled Trials Published in the *Chinese Medical Journal* From 2007 to 2008

Sun Jing,<sup>1</sup> Han Kun,<sup>2</sup> Qian Shou-chu,<sup>1</sup> and You Su-ning<sup>3</sup>

**Objective** This study aimed to assess the reporting quality of randomized controlled trials (RCTs) published in the *Chinese Medical Journal*.

**Design** According to the hand-search guidelines of Cochrane Collaboration, we hand-searched the RCTs published in the *Chinese Medical Journal* from 2007 to 2008; ultimately, 32 RCT reports were enrolled in this study. Based on the importance of the 22 items in the 2001 revised CONSORT checklist, we divided the 22 items into 25 evaluation items and added the other 2,

**Table 11. Assessment Items and Their Scores**

No.	Assessment Items	Total Scores	Quality Level
1	<b>Title and Abstract</b> How participants were allocated to interventions in title	7	Low
2	How participants were allocated to interventions in abstract	21	High
3	<b>Introduction</b> Scientific background and explanation of rationale	32	High
4	<b>Methods</b> Eligibility criteria for participants	23	High
5	The settings and locations where the data were collected	20	Medium
6	Precise details of the interventions intended for each group and how and when they were actually administered	31	High
7	Specific objectives and hypotheses	27	High
8	Clearly defined primary and secondary outcome measures	25	High
9	How sample size was determined	2	Low
10	Method used to generate the random allocation sequence	9	Low
11	Method used to implement the random allocation sequence (numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned	2	Low
12	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups	0	Low
13	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated	0	Low
14	Statistical methods used to compare groups for primary outcomes	32	High
15	<b>Results</b> Flow of participants through each stage (a diagram is strongly recommended). Describe protocol deviations from study as planned, together with reasons	9	Low
16	Dates defining the periods of recruitment and follow-up	6	Low
17	Baseline demographic and clinical characteristics of each group	28	High
18	Whether the analysis was by intention to treat	5	Low
19	For each primary and secondary outcome, a summary of results for each group	32	High
20	Estimate effect size and its precision (eg, 95% confidence interval)	2	Low
21	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses	4	Low
22	All important adverse events or side effects in each intervention group	14	Medium
23	<b>Discussion</b> Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes	32	High
24	Generalizability (external validity) of the trial findings	14	Medium
25	General interpretation of the results in the context of current evidence	32	High
26	<b>Others</b> Clinical trial registration	0	Low
27	Whether or not approved by the ethics committee and patients signed the informed consents	14	Medium

clinical trial registration and ethics, totaling 27 items (TABLE 11). Each RCT paper was evaluated using a 27-point table for quality assessment, and then we calculated the score for each RCT paper (full mark 27 points). To analyze the specific reasons for papers with low scores, we calculated the total score for each item of the 32 enrolled papers (full mark 32 points), respectively. We divided the 32 points into 3 categories according to the scores of papers: low quality (0-10 points), medium quality (11-20 points), and high quality (21-32 points).

**Results** The average score of articles obeyed the normal distribution with  $(13.2 \pm 2.6)$  points, less than the half of full mark (27 points). Ten items in the evaluation checklist belonged to the high-quality category (37.0%), 5 to the medium category (18.5%), and 12 to the low category (44.5%). None of the RCT reports described the randomization implementation, blinding or masking, and clinical trial registration. Only 2 reports (6.25%) mentioned how the sample size was determined and the method used to implement

the random allocation sequence. Nine reports (28.13%) described the method used to generate the random allocation sequence. Of the 32 items, the section of methods had the lowest score for poor description of sample size, randomization, and blinding.

**Conclusions** The reporting of RCTs published in the *Chinese Medical Journal* has not met the checklist of the CONSORT statement and needs to be improved. The RCT reporting should be specified in the Instructions for Authors of medical journals and be publicized among medical researchers, medical writers, readers, editors, and peer reviewers in China.

<sup>1</sup>*Chinese Medical Journal*, Chinese Medical Association, 42 Dongsi Xidajie, Beijing 100710, China, e-mail: sunj@cma.org.cn; <sup>2</sup>*Chinese Journal of Ophthalmology*, Chinese Medical Association, 42 Dongsi Xidajie, Beijing 100710, China; <sup>3</sup>Chinese Medical Association, 42 Dongsi Xidajie, Beijing 100710, China

### Application of Item Response Theory to Manuscript Rating Scales

Brian Budgell<sup>1</sup> and Takeo Nakayama<sup>2</sup>

**Objective** Content guidelines and rating scales derived from them, for example, CONSORT and the Jadad scale, are sometimes used to generate quantitative measures of manuscript quality. The initial inclusion of items is normally based on informed, expert opinion. However, the validation of scales and individual items can only be achieved following implementation. The objective of this presentation is to demonstrate the application of item response theory to the validation of checklist and rating scale items.

**Design** A corpus of 62 reports of randomized controlled trials (RCTs) in traditional medicine was analyzed for the prevalence of 24 selected CONSORT checklist items. Sensitivities of the checklist items were calculated as the ratio of prevalences of checklist items in the top and bottom quartiles of reports (based on summary scores). This permitted the post hoc generation of a rating scale with greater internal validity and sensitivity.

**Results** Prevalences of items in the 62 reports ranged from 0.08 (95% confidence interval [CI], 0.01-0.15) for item 11b (if done, how was the success of blinding evaluated) to 0.97 (95% CI, 0.93-1.00) for item 4a (precise details of the interventions intended for each group). Normalized summary scores for manuscripts ranged from 0.38 to 0.88 (mean, 0.63). Sensitivities (discriminatory indices for top versus bottom quartiles of reports) ranged from 1.07 to 15.00. Weighting checklist items according to prevalence and eliminating items with low sensitivities permitted the post hoc generation of a rating scale with greater internal validity and sensitivity.

**Conclusions** The validity and sensitivity of manuscript rating scales may be enhanced by the application of item response theory. The implementation of such validated scales would likely improve the editorial and review process and also inform decisions on the inclusion of reports in systematic reviews and meta-analyses.

<sup>1</sup>Canadian Memorial Chiropractic College, 6100 Leslie St, Toronto, Ontario M2H 3J1, Canada, e-mail: bs.budgell@gmail.com; <sup>2</sup>Department of Health Informatics, Kyoto University School of Public Health, Yoshida Konoe, Sakyo, Kyoto, Japan

### The Null Hypothesis Significance Test in Health Sciences Research (1995-2006): Statistical Analysis and Interpretation

Luis Silva,<sup>1</sup> Patricio Suarez,<sup>2</sup> Ana Fernandez,<sup>3</sup> Vanesa Alvarez,<sup>4</sup> and Tania Iglesias<sup>5</sup>

**Background** The null hypothesis significance test (NHST) is the most frequently used statistical method, although its inferential validity has been widely criticized since its introduction. In 1988, the International Committee of Medical Journal Editors (ICMJE) warned against sole reliance on NHST to substantiate study conclusions and suggested supplementary use of confidence intervals (CIs). Little research has examined the use of these statistical methods in the light of the ICMJE recommendation. The objective of this study was to evaluate patterns since 1995 in use of NHST and CIs both in English- and Spanish-language biomedical publi-

cations with particular focus on accuracy regarding interpretation of statistical significance and conclusion validity.

**Design** Original articles from 3 English and 3 Spanish biomedical journals in 3 fields (general medicine, clinical specialties, and epidemiology/public health) were considered for this study. Papers published in 1995-1996, 2000-2001, and 2005-2006 were selected through a systematic sampling method. After excluding the purely descriptive and theoretical articles, quantitative studies were evaluated for their use of NHST with *P* values and/or CIs for interpretation of statistical significance and relevance in study conclusions.

**Results** Among 1043 original papers, 874 were selected for review. The exclusive use of *P* values was less frequent in English-language publications as well as public health journals; overall such use decreased from 41.3% in 1995-1996 to 21.2% in 2005-2006. While the use of CIs increased over time, the fallacy of significance (to homologate statistical and substantive significance) appeared very often, mainly in journals devoted to clinical specialties (81.3%). On papers originally written in English and Spanish, 14.6% and 9.6% mentioned statistical significance in their conclusions, respectively.

**Conclusions** Although the exclusive use of NHST decreased over time in publications reviewed, this predominant pattern of statistical analysis remains slow to change. The communication of statistical results in the clinical setting, particularly among publications in Spanish, still presents considerable deficiencies.

<sup>1</sup>National Center for Medical Sciences Information Research, 27th, 110, Havana, 10400, Cuba, e-mail: lcsilva@infomed.sld.cu; <sup>2</sup>Cabueñes Hospital, Gijón, Spain; <sup>3</sup>Oviedo University, Oviedo, Spain; <sup>4</sup>Secondary School, Oviedo, Oviedo, Spain; <sup>5</sup>TeleCable of Asturias, Asturias, Spain

### Retractions

#### Why and How Do Journals Retract Articles?

Elizabeth Wager<sup>1</sup> and Peter Williams<sup>2</sup>

**Objective** Cases submitted to the Committee On Publication Ethics (COPE) suggest that journals and publishers do not have consistent policies about when and how to retract articles. As the first stage in developing retraction guidelines, we investigated why and how journals retract articles and editors' experiences of the process.

**Design** Analysis of retractions in PubMed from 1988 to 2008 with English text. We obtained all retractions for 2005-2008 and a random sample from 1988 to 2005 from journals available at University College London. Both authors extracted data and achieved consensus on classification. A purposive sample of editors was interviewed to learn about their experiences of retractions.

**Results** We analyzed 312 of the 529 retractions included in PubMed from 1988 to 2008 and interviewed 5 editors about 7 cases. Articles were retracted because of data fabrication (5%), data falsification (4%), plagiarism (16%), redundant publication

(17%), disputed authorship/data ownership (5%), inaccurate/misleading reporting (4%), honest research errors (28%), non-replicable findings (11%), or other/no stated reason (9%). Some journals also banned authors of plagiarized or redundant publications. Many retractions were issued by all or some authors (63%) but a significant proportion were issued by editors/publishers (29%) or others (8%). During interviews, editors described the considerable difficulties and significant workload in retracting articles when authors are uncooperative. Most retractions (87%) were of full papers reporting primary data but 13% were other article types (eg, literature reviews or letters). The retracted publications covered basic biomedical research (58%), clinical medicine (23%), and other subjects (19%) reflecting the composition of PubMed.

**Conclusions** Analysis of PubMed retractions combined with experience at COPE and published cases where journals have not retracted fraudulent articles indicates a considerable diversity of approach regarding how and why articles are retracted and sanctions imposed by journals for misconduct. Interviews suggested editors would welcome more guidance.

<sup>1</sup>Sideview, 19 Station Rd, Princes Risborough, Bucks, HP27 9DE, UK e-mail: liz@sideview.demon.co.uk; <sup>2</sup>University College, London, UK

### Round Up the Usual Suspects? Involvement of Medical Writers and the Pharmaceutical Industry in Retracted Publications

Karen L. Woolley,<sup>1,2</sup> Mark J. Woolley,<sup>2</sup> Rebecca A. Lew,<sup>2</sup> Narelle J. Bramich,<sup>2</sup> Julie A. Ely,<sup>2</sup> Serina Stretton,<sup>2</sup> Julie A. Monk,<sup>2</sup> and Janelle R. Keys<sup>2</sup>

**Objectives** (1) To quantify, for the first time, how involved declared medical writers and the pharmaceutical industry have been in publications retracted for misconduct; (2) to investigate factors associated with misconduct retractions.

**Design** We used PubMed (limits: English, human, January 1966-February 2008) to identify publications retracted for either misconduct or mistake. Standardized definitions and data collection tools were used (interrater reliability = 100%), and the mistake retractions served as the control group. Data were analyzed by an independent academic statistician.

**Results** Of the 463 retractions retrieved, 213 (46%) were misconduct retractions. The involvement of declared medical writers or the pharmaceutical industry was very low or nonexistent for misconduct retractions. Compared with mistake retractions, misconduct retractions were significantly associated with absence of declared medical writer or pharmaceutical industry involvement, single authorship, first author having at least 1 other retraction, or an affiliation in a low/middle-income country (TABLE 12).

**Conclusions** The involvement of declared medical writers or the pharmaceutical industry in misconduct retractions is very low or nonexistent. Our data challenge popular opinion and justify increased attention on factors that are associated with misconduct retractions.

**Table 12. Involvement in Retracted Publications**

Involvement in Retracted Publication	All Retractions, N = 463	Misconduct Retractions, n = 213	Odds Ratio of Retraction for Misconduct vs Mistake (95% CI)
Medical writer and pharmaceutical industry	2 (0.43%)	0 (0.00%)	NC
Medical writer	23 (4.97%)	3 (1.41%)	0.16 (0.05, 0.57)
Pharmaceutical industry	36 (7.78)	8 (3.76%)	0.25 (0.11, 0.58)
No pharmaceutical industry	427 (92.22%)	205 (96.24%)	3.74 (1.66, 8.40)
Single author	44 (9.50%)	26 (12.21%)	2.04 (1.01, 4.12)
First author with at least one other retraction	165 (35.64%)	94 (44.13%)	2.05 (1.35, 3.11)
First author affiliated with low/middle-income country	46 (9.94%)	28 (13.15%)	2.34 (1.18, 4.63)

CI, confidence interval; NC, noncalculable because of zero misconduct retractions.

<sup>1</sup>University of Queensland, Brisbane St Lucia, Australia, and University of the Sunshine Coast, Queensland, Australia; <sup>2</sup>ProScribe Medical Communications, 18 Shipyard Circuit, Noosaville Queensland 4566, Australia, e-mail: kw@proscribe.com.au

### Training

#### A Survey of Past Participants in the *Annals of Emergency Medicine* Editorial Board Fellowship Program

Teri Reynolds<sup>1</sup> and Michael Callahan<sup>2</sup>

**Objective** While initiatives to increase the number of physician researchers have been widespread, little has been published on the recruitment and training of physician editors early in their careers. *Annals of Emergency Medicine* established the Resident Editorial Fellow program in 1998 with 1 to 3 senior emergency medicine residents selected annually in a competitive process. These resident “fellows” serve during their penultimate or final year of residency and participate in editorial meetings, review manuscript development, shadow decision correspondence, complete article reviews, and oversee article selection for the resident section. We profiled the program by surveying past participants on their subsequent positions and the perceived impact of the fellowship.

**Design** In early 2009 we emailed a 7-question survey to all 14 prior fellows and asked them submit a curriculum vitae. The survey was self-directed and required free text responses.

**Results** A total of 14/14 completed the survey. Seven (50%) reported some editorial experience prior to the fellowship, largely on newsletters or nonscientific publications; 12 (86%) reported prior research. Nine (64%) currently serve on the editorial board, appointment to which is based on editorial performance and is not part of the fellowship. Thirteen (93%) currently review for at least 1 scientific journal and 7 (50%) for more than 2. In addition, as reviewers at *Annals of Emergency Medicine*, prior fellows on average are rated 4 (mean, 4.0, median, 4.2) out of possible

5, well above the average score for experienced reviewers. Ten (71%) hold full-time academic positions (2 recent fellows are still residents and not counted among these 10). Five (36%) now serve as their department's clinical research director, and 12 (86%) teach students or residents. Twelve (86%) identified mentored peer review, and 8 (57%) attending editorial board meetings as among the best parts of the program. Eight (57%) felt the fellowship was responsible for their current involvement in editorial work. Four (29%) felt that it increased their commitment to research. Fourteen (100%) report that the fellowship fully met or exceeded their expectations.

**Conclusions** Our results describe a young cohort who remain engaged in academic medicine and editorial process (as reviewers and editors) and who are in positions to mentor future editors and researchers. Respondents describe a high level of satisfaction with the program, and many remain involved with the editorial board.

<sup>1</sup>Highland Hospital Emergency Department, 1411 E 31st St, Oakland, CA 94602, USA, e-mail: reynoldst@gmail.com; <sup>2</sup>University of California, San Francisco, San Francisco, CA, USA

### Trial Registration

#### Association of Trial Registration With the Results and Conclusions of Published Trials of New Oncology Drugs

Nicolas Rasmussen,<sup>1</sup> Kirby Lee,<sup>2</sup> and Lisa Bero<sup>3</sup>

**Objective** To determine whether advance registration reduces bias against statistically insignificant results in the randomized controlled trial literature concerning new drugs.

**Design** This is a cross-sectional study of published reports of clinical trials evaluating the efficacy of drugs approved for new indications in oncology (where registration was first widely practiced) from 2000 through 2005. Relevant trial reports were identified using PubMed and the Cochrane Library. Evidence of trial registration in the year prior to publication was obtained by a search of public trial databases and corporate registries. Data on blinding, results for primary outcomes, and conclusions were extracted independently by 2 coders. Univariate and multivariate logistic regression identified associations between independent variables and favorable results and conclusions.

**Table 13. Association Between Characteristics of Articles and Statistically Significant Outcome or Conclusions That Favor the Test Drug: Multivariate Logistic Regression (N = 137)**

Characteristic	Category	Results Favor Test Drug			Conclusions Favor Test Drug		
		Favorable n/Total n (%)	OR (95% CI)	P	Favorable n/Total n (%)	OR (95% CI)	P
Trial registration before publication	No	44/83 (53)	1.00		55/83 (66)	1.00	
	Yes	36/54 (67)	1.50 (0.61-3.68)	.38	43/54 (80)	1.68 (0.64-4.43)	.29
Sponsored	No	14/28 (50)	1.00		15/28 (54)	1.00	
	Yes	66/109 (61)	1.15 (0.42-3.16)	.79	83/109 (76)	2.30 (0.87-6.07)	.09
Blinding	Not stringent	62/101 (61)	1.00		73/101 (72)	1.00	
	Stringent	18/36 (50)	0.42 (0.17-1.02)	.06	25/36 (69)	0.63 (0.25-1.59)	.33
Sample size	Quartile 1 (6-122)	10/34 (29)	1.00		19/34 (56)	1.00	
	Quartile 2 (123-352)	22/34 (65)	5.93 (2.00-17.57)	.62	24/34 (71)	2.28 (0.79-6.56)	.90
	Quartile 3 (353-772)	22/35 (63)	7.48 (2.26-24.7)	.25	26/35 (74)	2.77 (0.83-9.21)	.70
	Quartile 4 (773-8010)	26/34 (76)	14.07 (3.84-51.50)	.01	29/34 (85)	5.18 (1.32-20.28)	.08
Comparison group	Placebo	45/72 (63)	1.00		53/72 (74)	1.00	
	Active comparator only	35/65 (54)	0.48 (0.20-1.16)	.11	45/65 (69)	0.71 (0.29-1.74)	.46
Primary efficacy outcome	Surrogate	63/107 (59)	1.00		77/107 (72)	1.00	
	Survival	17/30 (57)	0.28 (0.09-0.84)	.02	21/30 (70)	0.36 (0.11-0.17)	.09

**Results** In univariate analyses, reports of trials unambiguously registered prior to publication (54/137) were more likely to describe statistically significant efficacy results and reach conclusions favoring the test drug (for results, odds ratio [OR], 1.77; 95% confidence interval [CI], 0.87-3.61). Reports of trials sponsored by the test drug maker and with larger sample sizes were significantly more likely to favor the test drug. In multivariate analysis, reports of prior registered trials again were not less likely to favor the test drug (for significant results: OR, 1.50; 95% CI, 0.61-3.68); larger sample sizes and surrogate outcome measures were statistically significant predictors of favorable results, while nonstringent blinding approached statistical significance. Subset analysis similarly showed that prior registered trial reports were more likely to describe statistically significant results favoring new drugs, among 109 industry-sponsored studies only (OR, 1.57; 95% CI, 0.60-4.11) and among 115 main reports only (ie, underlying trials) (OR, 1.30; 95% CI, 0.50-3.34). (TABLE 13.)

**Conclusions** Prior registration of trials alone did not decrease the trial literature's bias against statistically non-significant results. Additional mechanisms to ensure full reporting of trial results are necessary.

<sup>1</sup>National Drug and Alcohol Research Centre, University of New South Wales, Sydney NSW 2052, Australia, e-mail: N.Rasmussen@unsw.edu.au; <sup>2</sup>Department of Clinical Pharmacy, University of California, San Francisco, San Francisco, CA, USA; <sup>3</sup>Clinical Pharmacy and Health Policy, University of California, San Francisco, San Francisco, CA, USA

### Characterizing Sponsor-Imposed Restrictions on Disclosing Results of Clinical Trials

Tony Tse, Rebecca J. Williams, and Deborah A. Zarin

**Objective** Concern about undisclosed conflicts of interest and associated withholding of trial data by sponsors is growing. An FDA Amendments Act (FDAAA) provision mandating public disclosure of agreements that restrict the principal investigator's (PIs) ability to disclose results became effective on September 27, 2008. ClinicalTrials.gov includes the following categories that were based on published results of surveys of trial sponsors and organizations that conduct trials: short embargo ( $\leq 60$  days), no content control; longer embargo ( $>60$  days and  $\leq 180$  days); no content control; and other disclosure restrictions. The objective of this study was to characterize the types of sponsor-PI agreements reported to ClinicalTrials.gov and propose options for improving the existing categorization scheme.

**Design** Entries from all 183 results records posted at ClinicalTrials.gov (as of May 15, 2009) were evaluated, including all full-text descriptions of the Other category.

**Results** Of the 154 studies for which PIs were not employees of the sponsor, 117 (76%) indicated a restriction: 14 (12%) impose short embargoes, 14 (12%) impose longer embargoes, and 89 (76%) described Other restrictions. Among the 117 studies reporting restrictions, there were 33 phase 1-2 trials and 82 phase 3-4 trials; the majority (111/117) were sponsored by industry. Within the Other category, the following issues were addressed: (1) results communications for multisite studies (54/89); (2) "fixed" delays

after study completion, including for publication of multisite studies (40/89); (3) sponsors' rights to review, edit, and/or approve results communications (66/89); and (4) embargoes (80/89). Each sponsor used consistent text for its other entries. (TABLE 14.)

**Table 14. Sponsor-Imposed Restrictions Addressed in Other Category in Results Records Posted at ClinicalTrials.gov (as of May 15, 2009)**

Restriction Areas	No. of Trials (N = 89)	No. of Sponsors (N = 28)
<b>Multisite Studies</b>		
After multisite results disclosure	54	12
Time limit specified	38	3
No time limit specified	16	9
<b>"Fixed" Delays After Study Completion</b> (including those for multisite studies)		
12 months	26	5
18 months	10	3
24 months	4	3
<b>Control of Content</b>		
Sponsor review and comment	19	9
Sponsor can change confidential information only	39	9
Sponsor can change content	2	2
"Good faith"/"mutually agreeable" resolution of differences	3	3
Sponsor written consent/approval	3	3
<b>Embargoes</b>		
$\leq 60$ days	29	9
$>60$ and $\leq 180$ days	21	8
Unspecified amount of time	30	9

**Conclusions** Of the trials with sponsor-imposed restrictions on results disclosure, 76% were not captured by existing embargo categories at ClinicalTrials.gov. Our analysis suggests that additional categories would more accurately reflect common restrictions for multisite studies, fixed delays, sponsor control of content, and embargoes. Developing improved categories could enhance transparency by providing more consistent, comprehensive descriptions of PI-sponsor agreements.

National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894, USA, e-mail: atse@mail.nih.gov