



The Changing Landscape of Journalology in Medicine

Mitch Wilson, MD,^{*} and David Moher, PhD^{†,‡}

In the early 1970s, when *Seminars in Nuclear Medicine* started publication, little was known about the quality of reporting in biomedical journals. Senior scholars were invited to become scientific editors of journals based on their research credibility and stature. Their knowledge of journalology (publication science) was not assessed. Similarly, while the use of peer review was gaining momentum, there was limited guidance on the tasks and expectations of peer reviewing. Almost 50 years later, the evidence base regarding the quality of reporting is vast. This paper highlights some of this evidence including that relevant to imaging and nuclear medicine research. In biomedical publications, there is a crisis in reproducibility; high prevalence rates of reporting biases, such as selective outcome reporting; spin; low registration rates of research protocols; and endemic poor reporting of research across biomedicine. These issues and some more immediate solutions are also discussed in the paper. The use of reporting guidelines has been shown to be associated with better reporting of clinical trials and other research articles. The use of audit and feedback tools is likely to provide an important gauge about the functions of biomedical journals. Finally, the push to better equip scientific editors and peer reviewers is taking a more concerted effort.

Semin Nucl Med 49:105-114 Crown Copyright © 2018 Published by Elsevier Inc. All rights reserved.

Introduction

At the outset of the 1970s, when *Seminars in Nuclear Medicine* started publishing, authors interested in publishing research articles typically submitted a completed study to a journal of interest. Peer review of manuscripts was starting to gain momentum and many journals used this process to differentiate “inconsequential” submissions from “star” papers. Authors of accepted manuscripts typically signed over the intellectual copyright of their paper to the journal, acting on behalf of the publisher and/or society sponsoring the journal title. In the early 1970s, there was little data on the quality of reporting of published research. Schor and Karten reviewed statistical validity of conclusions

in 149 papers published during the first three months of 1964. Approximately two-thirds were deemed unacceptable although revision was possible, about a quarter (28%) were assigned “acceptable,” and the remaining 5% were judged “unsalvageable,” of no possible use.¹ Such analyses of the nuclear medicine literature had not yet been undertaken. Reporting guidelines had not yet been developed and there was little information on a host of issues about journalology (publication science).

Little was known about the effectiveness of peer review, the training of editors and peer reviewers, and the general operations of journals. Steven Lock, who assumed the editor’s mantle at the *BMJ* in 1975 (until 1991), was about to open up journals and push for evidence to inform their operations. He gave some indication of the difficulties of the job as part of his 1985 Rock Carling lecture entitled “A Delicate Balance: Editorial Peer Review in Medicine.” By 1975, Eugene Garfield had started calculating journal impact factors (JIFs), a metric of supposed journal importance. In 1986, Drummond Rennie, the deputy editor of *JAMA*, observed “there seems to be no study too fragmented, no hypothesis too trivial, no literature citation too biased or too

^{*}The Ottawa Hospital Research Institute, Ottawa, Canada.

[†]Centre for Journalology, Clinical Epidemiology Program, The Ottawa Hospital Research Institute, Ottawa, Canada.

[‡]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Canada.

Address reprint requests to David Moher, PhD, Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, The Ottawa Hospital – General Campus, 501 Smyth Rd, Room L1288, Ottawa, ON K1H 8L6, Canada. E-mail: dmoher@ohri.ca

egotistical, no design too warped, no methodology too bungled, no presentation of results too inaccurate, too obscure, and too contradictory, no analysis too self-serving, no argument too circular, no conclusions too trifling or too unjustified, and no grammar and syntax too offensive for a paper to end up in print.”² In an attempt to better understand peer reviewing and scientific editors, Rennie started the International Congress on Peer Review and Scientific Publication. The first meeting was held in 1989 in Chicago. It has been a successful venture in providing insight and evidence to help better understand peer review and the role and functions of scientific editors. The ninth congress will be held in September 2021.

Bad Reporting

Journals have existed for hundreds of years and are still the most important conduit for researchers to report the methods and results of their clinical and preclinical research. However, published articles are fraught with problems. Numerous reviews have identified widespread deficiencies in the reporting of research. Crucial aspects of study methods and results are frequently missing impeding transparency and reproducibility, essential parts of the research process.³ For instance, in 2012, the methods reporting and analysis strategies of 241 functional MRI articles were evaluated.⁴ Many of the publications did not report methodological details in sufficient detail for replication by independent researchers, and the proposed data analysis was highly variable across studies. An alarming number of researchers are publishing research in predatory journals making the studies difficult to identify.⁵ If reports of research are unusable or unidentifiable, this substantially reduces the ability to build on existing research. Poor reporting can also be misleading and decrease the confidence clinicians have in using evidence to inform practice, directly affecting patient care.^{6,7}

The reasons behind incomplete and misleading reporting of biomedical research are complex, involving multiple players. Scientific editors, peer reviewers, and researchers share responsibility. Some editors do not explicitly recommend the use of reporting guidelines as part of the review process⁸ despite emerging evidence indicating that their use is associated with more complete reporting.^{9,10} There appears to be a knowledge gap for scientific editors and peer reviewers affecting the value of their efforts and the trustworthiness of research publications¹¹; most editors and peer reviewers learn on the job without formal training. More than a third of manuscripts submitted to 46 journals were inappropriately classified as reports of randomized controlled trials (RCTs) by the editorial offices.¹² The inability of editors and peer reviewers to ensure scientific rigor is harmful and wasteful.

One estimate is that US\$240 billion is spent globally, annually, on health research.¹³ The outputs from this research are documented in about 3 million articles, of which about half are published in 25,000 journals (with a much larger number of editors). Eighty-five percent of this investment is avoidable waste (ie, it is modifiable). Manuscript peer review does not appear to prevent poor (and misleading) quality of reporting

of published research.¹⁴ Regardless of the reasons, it is important to remedy this situation. Globally, peer review is estimated to cost \$2.48 billion dollars annually and accounts for about one-quarter of the overall costs of scholarly publishing and distribution.¹⁵ The human cost was evaluated at 1.5 million hours every year.¹⁶ Overall, this is a regrettably poor return on a large fiscal and human investment.

Strategies to Improve the Accuracy, Completeness, and Transparency of Published Research

In 2014, in response to increasing concerns about waste in biomedical research, the Lancet published a “Waste in Research” series. In a forward to the series the editors asked, “our belief is that research funders, scientific societies, school and university teachers, professional medical associations, and scientific publishers (and their editors) can use this Series as an opportunity to examine more forensically why they are doing what they do...and whether they are getting the most value for the time and money invested in science.”¹⁷ There are several strategies available to help improve the quality of reporting biomedical research. Some of these are reviewed below. Any approach to improve reporting is more likely to succeed if it is endorsed, implemented, and audited.

Reporting Guidelines

While long-term solutions to improving the quality of reporting such as better training of scientific editors and peer reviewers are possible, more immediate solutions are available. Reporting guidelines, defined as a checklist of a minimal set of items to include when reporting a study, exist for a large number of study designs, including preclinical animal studies, randomized trials, observational studies, systematic reviews, and diagnostic accuracy studies. While there is accumulating evidence that use of reporting guidelines is associated with improved reporting though not in all cases,¹⁸ this evidence base is limited to only a few reporting guidelines.^{9,10} For journals, the most obvious problem they face is the quality of reporting of articles they publish. A systematic review involving 50 studies and reports of more than 16,000 randomized trials, assessing the effect of journal endorsement of the CONSORT checklist, showed improvements in the completeness of reporting for 22 checklist items.⁹

In 2003, STAndards for the Reporting of Diagnostic Accuracy studies (STARD), a reporting guideline that included a checklist of 25 items that should be reported in studies, was developed to increase the completeness and transparency of reporting of diagnostic accuracy studies.^{19,20} This was the first reporting guideline directly applicable to the radiological and nuclear medicine literature. Subsequent investigations have examined whether the quality of reporting in diagnostic accuracy studies has improved since the publication of STARD.²¹⁻²³ Smidt and colleagues examined the quality of

reporting in 12 high-impact-factor journals in 2000 (before STARD) and in 2004 (after STARD) using the STARD checklist.²¹ The authors reported that the mean number of reported STARD items increased from 11.9 (of 25) items in 2000 to 13.6 items in 2004. This represented a modest increase and the overall completeness remained suboptimal, with only slightly more than half of the STARD items being reported in 2004.

Korevaar and colleagues evaluated how well diagnostic accuracy studies published in 2012 adhered to the STARD guidelines and whether adherence had improved since 2000 and 2004.²² The mean number of STARD items reported in 2012 was 15.3 (of 25). An improvement of 3.4 items was observed compared with studies published in 2000 and an improvement of 1.7 items in the initial evaluation by Smidt in 2004. These results suggest that although some improvement has occurred since the creation of STARD, the overall quality of reporting in diagnostic accuracy studies was still suboptimal. Hong and colleagues were the first to specifically evaluate the adherence of diagnostic accuracy studies to STARD in imaging journals (including nuclear medicine journals).²³ Using the updated STARD checklist STARD 2015²⁴ consisting of 30 items, the authors demonstrated a mean adherence rate of 55% (16.6/30 items), representing a moderate adherence rate that is similar to rates documented in previous STARD adherence studies.^{21,22} It was also noted that studies published in higher impact journals and STARD adopter journals (journals that explicitly recommend adherence to STARD) reported more items.

While STARD is intended to help report individual diagnostic accuracy studies, the Preferred Reporting of Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA) is a 27-item checklist that provides guidance in the reporting of systematic reviews of diagnostic test accuracy studies. The PRISMA-DTA guideline can facilitate the transparent reporting of reviews. A recent assessment of the diagnostic accuracy literature by Salameh and colleagues evaluated the completeness of reporting of diagnostic test accuracy in systematic reviews from October 2017 to January 2018 using the PRISMA-DTA guidelines.²⁵ Of the 100 included reviews, the mean number of reported items from the PRISMA-DTA checklist was 18.6 (of 27), suggesting there is still room for improvement in the completeness of reporting of diagnostic accuracy systematic reviews.

The EQUATOR Network

The EQUATOR Network was established in 2006 and formally launched in 2008.²⁶ The vision was to develop a broad basket of tools to help authors, editors, peer reviewers, and others improve the reporting of articles published in biomedicine. Today, the network is on the way to meeting its initial remit. The EQUATOR library is an open repository of more than 400 reporting guidelines developed or currently under development.²⁷ The network also developed guidance to help others interested in developing a reporting guideline²⁸ and several toolkits for multiple stakeholders. These include guidance for authors writing manuscripts, manuscript peer

reviewers, and editors wanting to implement reporting guidelines at their journal. All four EQUATOR centers (Australia, Canada, France, and UK) have publication schools to help authors, particularly early career ones, produce better reports for publication consideration. The algorithm-based EQUATOR wizard is an initial attempt to help prospective authors identify the most appropriate reporting guideline to use when reporting their research.²⁹

There are now plenty of reporting guidelines to help authors, editors, and peer reviewers although several challenges remain. Whether there is reporting guideline inflation resulting in potential confusion for users requires consideration. Reporting guideline developers seem hesitant to provide data on the effectiveness of their developed reporting guideline. This might be related to the considerable problems in funding such endeavors. However, like pharmaceuticals, we should be more cautious about recommending the use of reporting guidelines without evidence of effectiveness. Even when armed with an initial evidence base about the effectiveness of reporting guidelines, few editors recommend that their peer reviewers use them.⁸ We need to enhance all implementation efforts.³⁰

While Lock had indicated the difficulties of running a journal, most of the journalology (publication science) literature focused on providing audit and feedback data about the quality of reporting from the authors' perspective. Little attention had been given to the role of the scientific editors (ie, editors are responsible for the content and policies of journals) and peer reviewers in this process. This is important because peer reviewers provide feedback to editors, who are the ultimate decision makers about the acceptability of research manuscripts.

Combating Reporting Biases Using Registration, Registered Reports, and Other Approaches

There are two broad categories of reporting biases. Nonreporting of complete studies (ie, publication bias), such as statistically negative reports of evaluations of a pharmaceutical agent is one category. Evidence of publication bias has existed from at least 1959.³¹ In 1986, Simes examined data contained in an oncology clinical trials registry and reported that statistically pooling the results of published trials only, compared with pooling published plus registered trials, provided clinicians and patients alike with differing and opposing estimates of the effectiveness of a cancer intervention.³²

The public health effects of publication bias have been recently exposed. Influenza antiviral medications such as oseltamivir (Tamiflu) have been commonly used and stockpiled by governments on the basis of recommendations by international organizations such as the World Health Organization.³³ Concerns have been raised, however, that these recommendations are supported by a body evidence affected by publication and reporting bias.³⁴ In 2014, these concerns prompted Jefferson and colleagues to conduct a systematic review of all available clinical study reports of RCTs examining

the benefits and harms of oseltamivir.³⁴ A total of 83 eligible clinical study reports were identified, of which 20 were included in the final analysis. It was demonstrated that in prophylactic studies, oseltamivir reduced the proportion of symptomatic influenza. In treatment studies, oseltamivir modestly reduced the time to first alleviation of symptoms but caused nausea and vomiting and increased the risk of headaches and renal and psychiatric complications. The disparity of findings between this systematic review, and those of the published research which promoted the widespread use of oseltamivir, highlights an important example of publication bias and its consequences. The findings of this systematic review provide compelling reason to question the recommendations that have led to the stockpiling of oseltamivir.

Sharifabdi and colleagues provide evidence for publication bias in imaging research.³⁵ They set out to determine whether higher reported accuracy estimates are associated with faster time to publication for imaging diagnostic accuracy studies. The authors measured diagnostic accuracy using Youden's index (sensitivity + specificity - 1, a single measure of diagnostic accuracy). With 55 systematic reviews and 781 primary studies included, they observed that Youden's index was negatively correlated with time to publication, suggesting a weak association between higher accuracy estimates and faster time to publication.

Another type of reporting bias is poorly describing reported characteristics. Here, an outcome identified in a protocol as of primary interest is demoted to one of secondary interest, without attribution, in the published completed study (eg, selective outcome reporting bias). A systematic review of the evidence for publication bias and selective outcome reporting bias in RCTs was conducted by Dwan and colleagues.³⁶ The review comprised of 16 cohort studies that have assessed publication bias and outcome reporting bias in RCTs. Eleven of the studies investigated publication bias and five investigated outcome reporting bias. The results of the systematic review reveal that both publication bias and selective outcome reporting bias are prevalent among RCTs. Of the 16 studies examined, 12 provided evidence that studies reporting positive or significant outcomes are more likely to be published. Three of the studies demonstrated that statistically significant outcomes were more likely to be reported than nonsignificant outcomes and that 40%-62% of studies had at least one primary outcome that was changed, added, or omitted in the publication, as compared with the study protocol.

A useful solution to countering reporting biases is to require registration of all research protocols. Although trial registration will not eliminate publication bias, it is likely to provide a degree of transparency and accountability not previously seen. While there are multiple trial registries, *clinicaltrials.gov* is by far the largest one having amassed 287,619 registries as of October 21, 2018, including an average of 86 new daily registrations in 2018.³⁷ While this is encouraging, it includes only a small fraction of all trials published. Efforts to bolster incentives to register trials and other research protocols are urgently needed. Academic promotion and tenure committees could include study registration as a criterion for

career advancement.³⁸ Indeed, formal registries for other designs still do not exist although there are various ways to register any type of study, such as Open Science Framework.

Other efforts are underway to reduce reporting biases. Several journals have initiated registered reports, first implemented in *Cortex* in 2013.³⁹ Here, authors submit their protocol, prior to initiating any data collection, to a journal for assessment and peer review. If the journal makes a positive assessment of the methods proposed and if the content area and methods meet a threshold, the journal makes an in-principle commitment to publishing the completed study regardless of the statistical results. Currently, 91 journals, included 21 in medicine, have adopted registered reports.⁴⁰ The Center for Open Science has also initiated a "preregistration challenge" in which 1000 researchers who register their design and methods (protocol) prior to data collection and publish their completed research prior to the end of 2018 are eligible for a \$1000 prize. Such efforts will also hopefully enhance efforts to tackle the crisis in reproducibility. Finally, the restoring invisible and abandoned trials (RIAT) initiative is another related effort.⁴¹ The RIAT initiative addresses the problem of publication bias from unpublished trials by providing the opportunity for third parties to access the underlying data of unpublished abandoned trials. These third parties, referred to as restorative authors, are able to analyze the data using the original trial protocol and then submit a manuscript for publication.

Spin

Spin has been defined as "specific reporting strategies, whatever their motive, to highlight that the experimental treatment is beneficial, despite a statistically nonsignificant difference for the primary outcome, or to distract the reader from statistically nonsignificant results."⁴² Spin is prevalent in the scientific literature including diagnostic studies.⁴³ A total of 126 diagnostic accuracy studies were examined, of which 53 evaluated the accuracy of imaging tests. The majority of studies were cross sectional (102) and longitudinal.¹⁷ To measure overinterpretation, two authors independently scored each article using a pretested data-extraction form which identified types of actual and potential overinterpretation. Of the 126 included articles, 39 (31%) contained a form of actual overinterpretation, including 29 (23%) with an overly optimistic abstract, 10 (8%) with a discrepancy between study aim and conclusion, and 8 (6%) with conclusions based on selected subgroups. In the analysis of potential overinterpretation, 89% of studies did not include a sample size calculation, 88% did not state a hypothesis, and 57% did not report confidence intervals for the diagnostic accuracy measurements. Of the 53 imaging studies, 16 (30%) contained forms of actual overinterpretation and 53 (100%) contained forms of potential overinterpretation. From these data, it is clear that overinterpretation of results is frequent in primary studies of diagnostic accuracy.

McGrath and colleagues assessed the frequency of overinterpretation in systematic reviews of diagnostic accuracy.⁴⁴ Overinterpretation was measured using a list of 10 items that represent actual overinterpretation in the abstract and/or

full-text of the review as well as a list of 9 items that represent potential overinterpretation. Of the 112 systematic review included, 40 (36%) assessed the diagnostic accuracy of imaging tests. The majority of reviews had a positive conclusion regarding the diagnostic accuracy of the test in the abstract (74%) and in the full-text (74%). There was at least one form of actual overinterpretation in 72% of the abstracts and in 69% of the full-texts. The most common forms of actual overinterpretation were a “positive conclusion, not reflecting the reported summary accuracy estimates,” occurring in 49% of abstracts and 50% of full-texts, and a “positive conclusion, not taking high risk of bias and/or applicability concerns into account” which occurred in 42% of abstracts and 23% of full-texts. Of the 112 reviews, 107 (96%) contained a form of potential overinterpretation, with “nonrecommended statistical methods for metaanalysis performed” most frequently identified (51%). The high prevalence of overinterpretation in systematic reviews of diagnostic accuracy identified by the authors may lead to optimism regarding test performance that is not justified. This in turn could result in errors in clinical decision-making and unnecessary increases in health care costs.

While our understanding of spin is increasing, particularly with more research, it is clear that editors and peer reviewers need to pay more attention to it, and its many manifestations. Similarly, it is now important to develop and evaluate interventions to detect and minimize the effects of spin in publications. Finally, whether patients and others reading articles make decisions based on spin needs investigation.

Deceptive Journals

Within the last decade, a new phenomenon has started to penetrate scientific publishing, particularly open access journals and publishers. These deceptive entities are often termed probable predatory journals (and publishers). Such journals typically compete with legitimate open access journals. They promise authenticity and scientific rigor, fast peer review, low author processing charges (publication fees), and immediate publication. These are “smart” promises because they feed directly into the problems and pitfalls of the majority of scientific journals. However, most of the evidence indicates that these journals do not use peer review, have fake journal metrics, including JIFs, and often do not meet their open access remit, asking authors to sign over copyright to the publisher.

Within a short period of time, articles in predatory journals have mushroomed from approximately 50,000 in 2010 to more than 400,000 in 2014.⁵ OMICS publishing group has been recognized as a source of predatory journals; they have approximately 700 journal titles. Nuclear medical imaging is not exempt from these publications. Predatory journals related to radiology/nuclear medicine published under OMICS include: *Journal of Nuclear Medicine and Radiation Therapy*, *Journal of Medical Physics and Applied Sciences*, *Imaging in Medicine*, and *Journal of Imaging and Interventional Radiology*.

While these numbers are getting larger, they are still a small fraction of the total number of articles published

annually. However, these articles and associated journals are starting to negatively impact what is considered trusted sources of medical knowledge, such as PubMed. Large fractions of certain content fields, such as neurology, have upwards of 25% predatory journals.

There are many opinions about predatory journals but not nearly enough research. Such research can improve an evidence base about predatory journal practices and help generate evidence-informed journal policy. It is unclear whether deceptive journals provide peer review. If they do, do authors find it substantive and helpful? Similarly, it is unclear why prospective authors submit to these journals. Are their motivations based on the author processing charge (APC) fees? Finally, are articles in predatory journals cited and included in systematic reviews?

Several authors and organizations have developed checklists and campaigns to help prospective authors avoid these deceptive journals. While such efforts are commendable, there is a risk that their advice may be discordant from one another, ultimately confusing the very audience the checklist developers are trying to help. Prospective authors might be most helped if a journal authenticator (internet browser plug-in) was developed to identify practice standards of all journals, not limited to deceptive publication practices journals, for prospective authors. Developing the journal authenticator fills a gap for authors, providing them with a set of core information about journal practices, such as the level of transparency in its operations, and if it is an open access journal, the creative commons license(s) it uses with authors.

What Does the Future Hold?

Incentivizing and Rewarding Authors for Better Behavior

Thus far our focus has been on the triad of authors, editors, and peer reviewers. While they are critical to helping improve the quality of publications, academic institutions are also very important in solving the problems of inadequate reporting of biomedical research; they are the intellectual home of authors and most often editors and peer reviewers. Universities are advancing the careers of the authors of these papers. Universities are still very much married to the JIF as part of assessing the merits of their faculty members. Some universities also give a fraction of a faculty member's salary to merit pay (ie, for publishing in journals with very high JIF). It has been reported that Chinese faculty receive tens of thousands of dollars for publishing in journals with very high impact factors.⁴⁵ Universities implicitly equate high JIF with quality. There is no corpus of data supporting the relationship between high JIF and the quality of articles it publishes. The JIF is a metric about the journal and provides no information about the quality of articles in the journal. It is important to disentangle any preconceived belief that these characteristics are correlated in a meaningful way.

If the quality of reporting is so poor, it makes the usability of the article limited and of questionable societal value.

Universities should consider whether their promotion and tenure criteria are fit for use in the 21st century. Promotion and tenure committees could easily modify their reward criteria away from “publish or perish” to include more evidence-based incentives, such as requiring authors to register their studies at inception and share their data once the studies are completed. Similarly, faculty could be rewarded for formal education and certification of peer review. Cumulatively these activities should play a more prominent role in rewarding faculty as part of their career progression. Such criteria also have considerable societal value.

Moving Away From Journal Impact Factors

The most prevalent metric used for assessing a journal's stature is examining its impact factor. A journal's impact factor is the average rate of citation to a given journal over a short time period. It is calculated by dividing the number of citations to any articles published in the journal over the past two years by the total number of research articles appearing in the journal over this period. There is no body of evidence associating quality (not to be confused with a journal's perceived prestige) of publications with JIF. Emphasis on JIF is irrational, since only 10%-20% of the papers published in a journal are responsible for 80%-90% of the citation impact of the journal, that is, its JIF.⁴⁶ There is a growing momentum away from using JIF as a criterion for assessing scientists.⁴⁷ The Declaration of Research Assessment (DORA) is a movement in this direction; scientists should not be assessed based on JIF of their publications.⁴⁸ Hundreds of organizations have signed DORA, including all seven UK research funders. Similarly, thousands of individuals have signed DORA. Similarly, there is a movement across journals to downplay their JIF. The JIF does have an important attribute; it can be quickly and easily ascertained. This is likely an important consideration if there is interest in substituting or adding more relevant evidence-based criteria to the assessments of scientists. Importantly, JIF does not measure the quality of an article. It measures something about the number citations of the journal in which the article was published.

Recently, Rosenkrantz and colleagues assessed impact in radiology journals using an alternative measure to impact factor/citation count.⁴⁹ The aim of the paper was to compare traditional citation count with an alternative impact metric, the Altmetric Attention Score (Altmetric). Altmetric tracks the number of weighted mentions that an article receives on a variety of online platforms, some of which include news posts, blog posts, Mendeley, Wikipedia, Twitter, and Facebook. Altmetric therefore represents a measure of the real-time online impact of an article. All 892 original investigations published in the 2013 issues of *Academic Radiology*, *American Journal of Roentgenology*, *Journal of the American College of Radiology*, and *Radiology* were included. The mean number of traditional citation counts was 10.7 ± 15.4 and the mean Altmetric score was 3.3 ± 13.3 . Among the articles, 96% had ≥ 1 traditional citation count while 42% of articles had an Altmetric score ≥ 1 . Citations and Altmetric were weakly associated ($r = 0.20$). Altmetric scores were higher in

articles with nonimaging content versus imaging content (5.1 ± 11.1 vs 2.8 ± 13.7 , $P = 0.006$). The authors concluded that although the overall online attention to radiology articles is low, the Altmetric score demonstrated unique trends, particularly for articles that are not imaging related, and may complement traditional citation counts as a measure of the impact of radiology journal articles.

Open Science and Open Access

There is now a movement toward open science, defined as the sharing of all information related to research. Open science can occur at all levels of research, most typically an individual research project. Here, a research team share the underlying data associated with the research project along with any methods, materials, and code associated with its conduct and results. In addition to making open any registration, protocol and findings in publications, open science will be a key driver of ensuring reproducibility. Patients appear supportive of data sharing related to studies in which they have participated.⁵⁰ It is also possible to have open science at a group level, such as a laboratory, where electronic laboratory books (eg, iPads) are openly shared. Finally, open science at the institutional level is possible, where there is no patent protection to discovery.⁵¹

Data sharing policies across journals need improvement to keep pace with the open science movement. To assess the prevalence of data sharing policies in the biomedical literature, Vasilevsky and colleagues reviewed the data sharing requirements of 318 biomedical journals.⁵² They reported that 11.9% of journals explicitly stated that data sharing was required for publication, 9.1% of journals required data sharing but did not state it would affect publication, and 23.3% of journals encouraged authors to share data but did not require it. Journals with data sharing policies facilitate reproducibility efforts.⁵³

Open access is a movement toward openness. Broadly speaking, there are two publication models for disseminating biomedical research. First, the traditional subscription model is based on authors submitting manuscripts to a journal for publication consideration and if accepted they transfer the copyright of the article's content to the journal/publisher. Interested readers are required to pay a fee to access the work. Fees can be paid via individual article purchases, through individual journal subscriptions, or a third party, such as a university library subscription, typically as part of a bulk purchase from a publisher. This publication model poses a barrier for interested readers and/or institutions without sufficient fiscal resources. Further, with this model, the copyright of the work is transferred to the journal/publisher which creates a barrier for additional future use by the authors and others, and limits knowledge translation activities.

Second, the open access model was established approximately 20 years ago to combat the inherent inequity of the long-standing subscription model. Here, if an article is accepted for publication, the author retains the copyright, meaning that the author and others can freely build upon and use the article's content. Authors of accepted articles are

charged an APC. This fee was meant to cover cost of internal journal activities, such as copy editing, ensure the journal's contents remained available in perpetuity, and that its contents are freely available to all, regardless of their fiscal abilities. This "levelling of the field" has a number of advantages, is morally uplifting, and has attracted considerable traction. There are a growing number of highly respected and influential open access journals and publishers. A downside of this model is the APC, which can range from \$2500 to \$5000, particularly for biomedical journals.

Narayan and colleagues recently assessed the extent to which open access policies have been adopted by radiology journals.⁵⁴ A total of 49 radiology journals (impact factors > 1.0) were included in the analysis, although nuclear medicine journals were excluded. Of the 49 journals, 36 (73%) had an open access option and 4 (8%) were exclusively open access. Open access policies have therefore been adopted by the majority of radiology journals, with a few journals being exclusively open access. Seminars in Nuclear Medicine offer a hybrid model, whereby authors can pay an APC for gold open access.

Training Editors and Peer Reviewers

The importance of biomedical journals cannot be overstated. What journals publish impacts patient care. Given the large number of biomedical journals—one estimate is 25,000—it is likely there is a very wide variation in the competencies of scientific editors. If editors, like other professionals, want to ensure the standards of their respective journals' attention is needed to ensure they are trained to the highest possible standards. Many scientific editors of biomedical journals operate largely without formal training and universal certification is not yet a high priority.⁵⁵ Instead, editors generally are invited to serve in their role by publishers, based on their expertise and stature in the field, since such expertise is essential for evaluating research and stature is important for establishing the reputation of the journal and attracting submissions. However, such expertise does not guarantee that editors have the background or training necessary to carry out their editor roles and responsibilities. More than a third of manuscripts submitted to 46 journals were inappropriately classified as reports of RCTs by the editorial offices.¹² The inability of editors and peer reviewers to ensure scientific rigor is harmful and wasteful.

One group of researchers has focused on developing a set of core competencies for scientific editors.⁵⁶ They developed a set of core competencies resulting from a series of activities including a scoping review, a Delphi process, and a face-to-face consensus meeting. Multiple relevant players, representing various groups, were involved throughout the process. The 14 key core competencies were divided into 3 major areas (editor qualities and skills, such as exercising sound judgment in making editorial decisions; publication ethics and research integrity, such as identify and assess problems related to selective reporting of publications, outcomes, and analyses, discussed in further detail below; and editorial principles and processes, such as interpreting journal and

scholarly metrics and ensure that these metrics are not manipulated in a way that is unfair or unscrupulous). Each competency has a list of associated elements or descriptions of more specific knowledge, skills, and characteristics that contribute to its fulfillment. These competencies are gaining traction and endorsement throughout the scholarly editorial community. What is needed now is an experimental evaluation assessing whether exposure to these core competencies produces more knowledgeable editors and more complete, accurate, and transparent reports of biomedical research.

Globally, peer review is estimated to cost \$2.48 billion dollars annually and accounts for about one-quarter of the overall costs of scholarly publishing and distribution.¹⁵ The human cost was evaluated at 15 million hours every year.¹⁶ Unfortunately, there is little by way of core competencies for peer reviewers, a much larger community than scientific editors. Editors often rely on peer reviews to help decide about publication acceptability. However, the evidence from two systematic reviews suggests that if peer review is effective, the effect is minimal at best.^{14,57}

While there are several commercial short courses available, very little is freely available for early career researchers wanting to learn more about the skills needed for conducting effective peer review. To help fill this gap, Publons, a commercial group within the Clarivate Analytics umbrella, has created a free 10-module Publons Academy course to provide training opportunities for peer reviewers, particularly early career ones. Part of the program requires the student to identify a mentor, often their supervisor, to help them complete two reviews. Once the "student" has completed the program, they receive a certificate.

Preprint Servers

Ever since *Philosophical Transactions* was established more than 300 years ago, journals have had difficulties maintaining efficiency. In one classic case, *Philosophical Transactions* lost a submitted paper for several years. There are several reasons why these problems exist but two reasons stand out. First, there is considerable peer reviewer fatigue, partly because this activity is largely carried out without remuneration and real credit, and requests to peer review keeps increasing. Peer review activities are carried out usually on evenings and weekends as additional contributions to science, once the reviewer's paid work has been completed. While a few journals have full time paid editors, most journals operate with part time editors who are paid a small stipend for their activities. Publishing issues have become far more complicated over time making the jobs of editors, particularly decision-making ones (ie, scientific editors), far more complicated. This can create a tremendous delay in the process of making decisions about the acceptability of manuscripts for publication, often locking up important and relevant research for many months or, at times, years. These problems are particularly acute for researchers, particularly early career ones.

Even when published, these publications may not be accessible to interested readers. It requires open access to content. To help overcome these issues, and others, scientists

have sought ways to share their research. One such mechanism has been the development of preprint servers, defined as “a scholarly manuscript posted by the author(s) in an openly accessible platform, usually before or in parallel with the peer review process.”⁵⁸ One of the earliest preprint servers was Arxiv, developed in the early 1990s by physicists for sharing papers, although informal preprint servers have existed for about 50 years.⁵⁹

Preprints allow researchers to “claim” some ownership of a research theme/project; this can be particularly useful for early career researchers in a highly competitive research environment. Preprint servers provide digital object identifier for each included manuscript. This information can be included in grant applications. Indeed, progressive granting agencies are recommending applicants include preprints in their CVs (eg, National Institutes of Health, USA). There are already several of them, including those for the biosciences (eg, BioRxiv) which have epidemiology and clinical trials subcategories. Many more preprint servers are coming to fruition, including those for medicine (Medrxiv). This has led to the formation of the scientific-driven initiative ASAPbio (Accelerating Science and Publication in biology) to promote their use.⁶⁰ Some journals are joining these efforts. For example, PLOS Medicine has partnered with BioRxiv and The Lancet has created a preprint server hosted by SSRN (previously known as Social Science Research Network).

It is possible that preprints will change the publication landscape. Instead of authors submitting a manuscript to a journal for publication consideration, preprint servers may turn this approach on its head.⁶¹ Here, authors will submit their manuscript to a preprint server. Journals will regularly scan these servers, perhaps using artificial intelligence algorithms, to identify papers of potential interest to their readers. In such cases, a journal editor will reach out to the preprint authors and invite them to formally submit their manuscript to the journal for publication consideration.

Audit and Feedback

Manufacturing industries have a long history of quality control for good reason. Airline manufacturers do not want planes falling out of the sky and steel makers do not want buildings collapsing. Quality control is typically carried out by providing an audit of the quality of the manufacturing processes and feeding this information back to the responsible personnel. Any deficiencies are discussed and, if necessary, modifications are made to the manufacturing processes to ensure the highest possible quality of the product.

Biomedical journals also produce a very important product, namely, research articles. However, there appears to be little by way of auditing the product and this information is not typically fed back to authors or readers of the journal. Auditing the quality of reports a journal publishes could provide important baseline information for editors and peer reviewers and identify problems as well as opportunities to enhance the published product. Making such information available to the public would send a strong positive signal about openness, sharing data and the journal’s commitment

to continuous quality improvement. There is a large volume of evidence indicating the benefits of audit and feedback.⁶²

The EBM Data lab, University of Oxford, has a number of excellent projects where auditing plays an important role. For example, the COMPare Trials project initially provided an audit of outcome reporting bias in five well-known journals that published 67 randomized trials over a 6-week interval in 2015.⁶³ The TrialsTracker project provides feedback about registration of randomized trials across a very broad range of academic institutions.⁶⁴ An institution conducting at least 30 trials annually can type their name into the program and be provided with feedback as to the fraction of the institution’s trials registered. Journals could also use this information and create additional information as part of any audit and feedback.

Contributor Role Taxonomy (CRediT) for Contributing to Research Output

Most journals, including *Seminars in Nuclear Medicine*, require a traditional form of authorship for submitted manuscripts. For example, to be an author requires some combination of tasks, such as conceptualizing the project, data analysis, and/or writing or reviewing draft versions. Authorship is laid out by in the International Committee of Journal Editors. When authorship guidance was first developed, there may have been a sound rationale for providing such guidance. Evidence suggests that a high fraction of authors do not meet the ICMJE’s criteria.⁶⁵ There is also evidence of authorship inflation⁶⁶ and manipulations of authorship such as guest authorship and ghost authorship.⁶⁷ In an attempt to rectify these problems and others, several proposals have been made.

Rennie and Yank proposed contributorship in an attempt to deal with more equitable contributions researchers make to a research project.⁶⁸ CRediT is an effort to take contributorship further.⁶⁹ It recognizes that contributions to a research project can be varied and not necessarily part of a team that conceptualized a research project. For example, CRediT developed a taxonomy of 14 possible roles, such as software management which might include “Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components developing specific code.” Similarly, with the complexity of data, another role is visualization: “Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.” There are also more traditional roles, including writing. As open science becomes more prevalent and the accepted way of research, there is likely a need to promote ways to integrate more diverse contributions. A CRediT byline might look something like “Conceptualization, S.C.P. and S.Y.W.; Methodology, A.B., S.C.P., and S.Y.W.; Investigation, M.E., A.N.V., N.A.V., S.C.P., and S.Y.W.; Writing – Original Draft, S.C.P. and S.Y.W.; Writing – Review & Editing, S.C.P. and S.Y.W.; Funding Acquisition, S.C.P. and S.Y.W.; Resources, M.E.V and C.K.B.; Supervision, A.B., N.L.W., and A.A.D. (Cell Press). A number of publishers endorse the CRediT system” <https://www.cell.com/pb/assets/raw/shared/guidelines/CRediT-taxonomy.pdf>.

References

- Schor S, Karten I: Statistical evaluation of medical journal manuscripts. *JAMA* 195:1123-1128, 1966
- Rennie D: Guarding the guardians: A conference on editorial peer review. *JAMA* 256:2391-2392, 1986
- Glasziou P, Altman DG, Bossuyt P, et al: Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 383:267-276, 2014
- Carp J: The secret lives of experiments: Methods reporting in the fMRI literature. *Neuroimage* 63:289-300, 2012
- Shen C, Björk BC: 'Predatory' open access: A longitudinal study of article volumes and market characteristics. *BMC Med* 13:230, 2015
- Duff JM, Leather H, Walden EO, et al: Adequacy of published oncology randomized controlled trials to provide therapeutic details needed for clinical application. *J Natl Cancer Inst* 102:702-705, 2010
- Dancey JE: From quality of publication to quality of care: Translating trials to practice. *J Natl Cancer Inst* 102:670-671, 2010
- Hirst A, Altman DG: Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals. *PLoS One* 7: e35621, 2012
- Turner L, Shamseer L, Altman DG, et al: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev* 2012(11). <https://doi.org/10.1002/14651858.MR000030.pub2>
- Stevens A, Shamseer L, Weinstein E, et al: Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: Systematic review. *BMJ* 348:g3804, 2014
- Galipeau J, Moher D, Campbell C, et al: A systematic review highlights a knowledge gap regarding the effectiveness of health-related training programs in journalology. *J Clin Epidemiol* 68:257-265, 2015
- Hopewell S, Boutron I, Altman DG, et al: Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: Results of a randomised controlled trial. *BMC Med* 14:199, 2016
- Chalmers I, Glasziou P: Avoidable waste in the production and reporting of research evidence. *Lancet* 374:86-89, 2009
- Bruce R, Chauvin A, Trinquart L, et al: Impact of interventions to improve the quality of peer review of biomedical journals: A systematic review and meta-analysis. *BMC Med* 14:85, 2016
- Peer Review in Scientific Publications. [Report] House of Commons Science and Technology Committee. Eighth Report of Session 2010–12. Last Accessed January 31, 2018
- Peer Review: How We Found 15 Million Hours of Lost Time. American Journal Experts. Durham, NC. Last Accessed January 31, 2018
- Kleinert S, Horton R: How should medical science change? *Lancet* 383:197-198, 2014
- Botos J: Reported use of reporting guidelines among JNCI: Journal of the National Cancer Institute authors, editorial outcomes, and reviewer ratings related to adherence to guidelines and clarity of presentation. *Res Integr Peer Rev* 3:7, 2018
- Bossuyt PM, Reitsma JB, Bruns DE, et al: The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Ann Intern Med* 138:W1-W12, 2003
- Bossuyt PM, Reitsma JB, Bruns DE, et al: Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Clin Chem Lab Med* 41:68-73, 2003
- Smidt N, Rutjes AW, Van der Windt DA, et al: The quality of diagnostic accuracy studies since the STARD statement Has it improved? *Neurology* 67:792-797, 2006
- Korevaar DA, Wang J, van Enst WA, et al: Reporting diagnostic accuracy studies: Some improvements after 10 years of STARD. *Radiology* 274:781-789, 2014
- Hong PJ, Korevaar DA, McGrath TA, et al: Reporting of imaging diagnostic accuracy studies with focus on MRI subgroup: Adherence to STARD 2015. *J Magn Reson Imaging* 47:523-544, 2018
- Bossuyt PM, Reitsma JB, Bruns DE, et al: STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *Clin Chem* 2015. <https://doi.org/10.1373/clinchem.2015.246280>
- Salameh JP, McInnes MD, Moher D, et al: Completeness of reporting of systematic reviews of diagnostic test accuracy based on the PRISMA-DTA reporting guideline. *Clin Chem* 2018. <https://doi.org/10.1373/clinchem.2018.292987>
- Thornton H: Report on EQUATOR network launch meeting 26th June 2008 "Achieving Transparency in Reporting Health Research". *Int J Surg* 6:428-431, 2008
- Enhancing the QUALity and Transparency Of health Research. <http://www.equator-network.org/> Accessed November 1, 2018.
- Moher D, Schulz KF, Simera I, et al: Guidance for developers of health research reporting guidelines. *PLoS Med* 7:e1000217, 2010
- The EQUATOR wizard: A new tool to help authors find the right reporting guideline. <http://www.equator-network.org/2015/12/17/findtherightreportingguideline/> Accessed November 1, 2018.
- Shamseer L, Hopewell S, Altman DG, et al: Update on the endorsement of CONSORT by high impact factor journals: A survey of journal "Instructions to Authors" in 2014. *Trials* 17:301, 2016
- Sterling TD: Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *J Am Stat Assoc* 54:30-34, 1959
- Simes RJ: Publication bias: The case for an international registry of clinical trials. *J Clin Oncol* 4:1529-1541, 1986
- World Health Organization. WHO model list of essential medicines, 20th list (March 2017, amended August 2017)
- Jefferson T, Jones M, Doshi P, et al: Oseltamivir for influenza in adults and children: systematic review of clinical study reports and summary of regulatory comments. *BMJ* 348:g2545, 2014
- Sharifabadi AD, Korevaar DA, McGrath TA, et al: Reporting bias in imaging: Higher accuracy is linked to faster publication. *Eur Radiol* 28(9):3632-3639, 2018
- Dwan K, Altman DG, Arnaiz JA, et al: Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One* 3:e3081, 2008
- ClinicalTrials.gov: Trends, Charts, and Maps. <https://clinicaltrials.gov/ct2/resources/trends>. Accessed October 16, 2018
- Moher D, Naudet F, Cristea IA, et al: Assessing scientists for hiring, promotion, and tenure. *PLoS Biol* 16:e2004089, 2018
- Chambers CD: Registered reports: A new publishing initiative at Cortex. *Cortex* 49:609-610, 2013
- Hardwicke TE, Ioannidis JPA: Mapping the universe of registered reports. *OSF* 2018 <https://osf.io/gjsaf/> Accessed November 1, 2018
- Doshi P, Shamseer L, Jones MA, et al: Restoring biomedical literature with RIAT. *BMJ* 361:k1742, 2018
- Boutron I, Dutton S, Ravaud P, et al: Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA* 303:2058-2064, 2010
- Ochodo EA, de Haan MC, Reitsma JB, et al: Overinterpretation and misreporting of diagnostic accuracy studies: Evidence of "spin". *Radiology* 267:581-588, 2013
- McGrath TA, McInnes MD, van Es N, et al: Overinterpretation of research findings: Evidence of "spin" in systematic reviews of diagnostic accuracy studies. *Clin Chem* 63(8):1353-1362, 2017
- Quan W, Chen B, Shu F: Publish or impoverish: An investigation of the monetary reward system of science in China (1999-2016). *Aslib J Inf Manag* 69:486-502, 2017
- Garfield E: The history and meaning of the journal impact factor. *JAMA* 295:90-93, 2006
- Benedictus R, Miedema F, Ferguson MW: Fewer numbers, better science. *Nat News* 538:453, 2016
- Curry S: Let's move beyond the rhetoric: It's time to change how we judge research. *Nature* 554:147, 2018
- Rosenkrantz AB, Ayooola A, Singh K, et al: Alternative metrics ("altmetrics") for assessing article impact in popular general radiology journals. *Acad Radiol* 24:891-897, 2017
- Mello MM, Lieou V, Goodman SN: Clinical trial participants' views of the risks and benefits of data sharing. *N Engl J Med* 378:2202-2211, 2018
- Rouleau G: Open science at an institutional level: An interview with Guy Rouleau. *Genome Biol* 18:14, 2017
- Vasilevsky NA, Minnier J, Haendel MA, et al: Reproducible and reusable research: Are journal data sharing policies meeting the mark? *Peer J* 5: e3208, 2017

53. Naudet F, Sakarovitch C, Janiaud P, et al: Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: Survey of studies published in The BMJ and PLOS Medicine. *BMJ* 360:k400, 2018
54. Narayan A, Lobner K, Fritz J: Open access journal policies: A systematic analysis of radiology journals. *J Am Coll Radiol* 15:237-242, 2018
55. Moher D, Altman DG: Four proposals to help improve the medical research literature. *PLoS Med* 12:e1001864, 2015
56. Moher D, Galipeau J, Alam S, et al: Core competencies for scientific editors of biomedical journals: Consensus statement. *BMC Med* 15:167, 2017
57. Jefferson T, Rudin M, Folsie SB, et al: Editorial peer review for improving the quality of reports of biomedical studies. *Cochrane Database Syst Rev* 2006 (1). <https://doi.org/10.1002/14651858.MR000016.pub2>
58. Committee on Publication Ethics. Discussion document on preprints. [ONLINE]. Available at: https://publicationethics.org/files/u7140/COPE_Preprints_Mar18.pdf [Accessed September 28, 2018]
59. Klein SR: On the origins of preprints. *Science* 358:602, 2017
60. ASAPbio. <http://asapbio.org/>. Accessed November 1, 2018
61. Berlin S: If the papers don't come to the journal. . . . *EMBO Rep* 2018: e201845911
62. Ivers N, Jamtvedt G, Flottorp S, et al: Audit and feedback: Effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 6 (6), 2012. <https://doi.org/10.1002/14651858.CD000259.pub3>
63. COMPare: Tracking switched outcomes in clinical trials. <http://compare-trials.org/>. Last Accessed October 16, 2018
64. Powell-Smith A, Goldacre B: The TrialsTracker: Automated ongoing monitoring of failure to share clinical trial results by all major companies and research institutions. *F1000Res*: 5, 2016
65. International Committee of Medical Journal Editors: Defining the Role of Authors and Contributors. <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html>. Accessed November 1, 2018
66. Tilak G, Prasad V, Jena AB: Authorship inflation in medical publications. *Inquiry* 52:0046958015598311, 2015
67. Moher D: Along with the privilege of authorship come important responsibilities. *BMC Med* 12:214, 2014
68. Rennie D, Yank V, Emanuel L: When authorship fails: A proposal to make contributors accountable. *JAMA* 278:579-585, 1997
69. Brand A, Allen L, Altman M, et al: Beyond authorship: Attribution, contribution, collaboration, and credit. *Learn Publ* 28:151-155, 2015