Assessment of research quality in major infertility journals

Demián Glujovsky, M.D., M.Sc., a,b Bárbara Riestra, M.D., b Andrea Coscia, M.D., b Carolina Boggino, M.D., b Daniel Comandé, B.Sc., a and Agustín Ciapponi, M.D., M.Sc. a

a Southern American Branch of the Iberoamerican Cochrane Centre, Institute for Clinical Effectiveness and Health Policy (IECS); and b Center for Studies in Gynecology and Reproduction (CEGYR), Buenos Aires, Argentina

Objective: To evaluate the level of evidence published in infertility journals with the highest impact factors.

Design: Systematic review. We searched in PubMed identifying potential systematic reviews with meta-analysis (SRs) and randomized controlled trials (RCTs) between 2006 and 2010 in the five fertility journals with highest impact factor.

Setting: Academic institution.

Patient(s): None.

Intervention(s): None.

Main Outcomes Measure(s): Number and proportion of potential SRs and RCTs published in 2006–2011, and quality assessment of real RCTs published in 2010.

Result(s): Among evaluated articles, 1.5% and 6.8% were SRs and RCTs. Fertility and Sterility has been the journal with more potential SRs and RCTs, and Human Reproduction Update was the only one with an increasing trend in the number of potential SRs (from 5.3% in 2006 to 24.4% in 2011). Among confirmed RCTs, for each quality assessment item, between 50% and 85% were classified as having low risk of bias, and the most common high risk of bias was related to allocation concealment. Only 23% had a trial registration, which were associated with higher quality assessment classifications. Only 10% of RCTs reported pharmaceutical industry funding.

Conclusion(s): This is the first study assessing the methodologic quality of publications in the top five fertility journals. More than 90% of all the publications were neither SRs nor RCTs. It is remarkable that the risk of bias was generally low. Enhancing trial registration and funding source statements represent opportunities to improve the quality of reporting. We hope this information is useful to researchers, editorial boards, and clinicians moving forward with research in our field. (Fertil Steril® 2012;98:1539–43. ©2012 by American Society for Reproductive Medicine.)

Key Words: Risk of bias, quality assessment, evidence, randomized controlled trials, systematic reviews

Q uestion of research and its publication is a fundamental aspect of medicine. In 2007, the British Medical Journal published a survey in which readers, editors, and advisers considered that evidence-based medicine should be one of the 15 most important medical milestones. It is widely accepted that the “pyramid of evidence” has systematic reviews (SRs) and randomized controlled trials (RCTs) at the top, especially for intervention effectiveness (1). An SR is a method of assessing systematically all the evidence available on a specific topic, using secondary data. Randomized controlled trials are considered the study types with the highest level of evidence, when using primary data. Of course, these types of studies do not answer all the clinical questions, but they constitute a good landmark to know where we stand and to establish the present trend of our publications.

Infertility is a topic that has been widely studied and published at the Cochrane Collaboration. The Menstrual Disorders and Subfertility Group is one of the three Cochrane groups with more SRs published in the collaboration (2). These topics have been extensively evaluated with SRs. A recent publication by Legro and Kunselman (3) showed that the term “meta-analysis” has been increasingly included in papers submitted to Fertility and Sterility. However, we ignore the proportion of RCTs published in peer-reviewed journals. Infertility is not an easy topic to study, because treatments are financially and emotionally costly; therefore, recruiting patients to participate in RCTs is usually difficult. On the other hand, multiple variables and potential confounders are usually present, suggesting that only good-quality studies will be able to produce confident
RESULTS

The five journals with the highest impact factor were the following: Human Reproduction Update (HR-Update), Human Reproduction (HR), Fertility and Sterility (F&S), Reproduction (Cambridge, England), and Reproductive BioMedicine Online (all official impact factors above 2.2). During the 6 years reviewed, 6.8% of all the publications were potential RCTs (a range of 132–185 per year). Two thirds of them were published by F&S and almost a quarter by HR [Fig. 1, Table 1]. Fertility and Sterility was the journal with the highest proportion of potential RCTs among all the publications (8.5% ± 1.3%), with more potential RCTs (114 ± 14 per year). When we did a subgroup analysis evaluating only the three journals that published mainly clinical studies with primary data, the average of potential RCTs slightly increased to 7.6%. During the same time period, 1.5% of all the publications were potential SRs (range, 25–48 per year). Almost half of them were published by F&S and a quarter by HR-Update [Fig. 2, Table 1]. Fertility and Sterility has been the journal with more potential SRs with meta-analysis in the last 6 years (17.1 ± 4.3 per year) and HR-Update the one with a trend toward increasing the number of potential SRs with meta-analysis (from 4 to 19 in the years studied); HR-Update showed the highest proportion of potential SRs among the published studies (with an increasing trend from 5.3% in 2006 to 24.4% in 2011). The rest of the journals published less than 2% of potential SRs.

In 2010, 152 of 169 potential RCTs retrieved and reviewed (89.9%) were confirmed as RCTs. When we analyzed the affiliation of the first author, the most represented countries were the United States (17.9%), Italy (10.6%), Turkey (7.2%), and Egypt (6.6%); of significance, 15.3% of the trials were multicentric. The topics most commonly studied were assisted reproductive technologies (21.9%), gonadotropins (8.9%), and artificial insemination (7.5%). Standard practice was the most common control arm (37%), whereas the rest were placebo (31.5%), other (20.5%), and no intervention (11%). Funding was disclosed by 39.7%, remained unclear in 51%, and 9.3% were not funded. Of those trials that were funded, 26% were sponsored by pharmaceutical companies. Conflict of interest was disclosed by 12.7% and remained unclear in
12.7%, whereas 74% claimed not to have conflicts of interests. Only 23% of the analyzed studies had a trial registration, mainly in ClinicalTrials.gov.

When analyzing the quality assessment of the RCTs, the most common risk of bias was related to allocation concealment and blinding of participants, which was classified as unclear or with high risk of bias in almost 50% of studies (Fig. 3). Low risk of bias for random sequence generation was 70%, and low risk for blinding of outcome assessment was 65%, whereas low risk for selective reporting and other biases was more than 85%. When doing a subgroup analysis by journal, no significant differences were found among them (data not shown, available on demand). Most of the trials had a simple parallel-groups design (91.9%), whereas a few used a cross-over design or clusters. The methodology for randomization most commonly used was a computer-generated list.

When we analyzed the risk of bias according to the trial registration status, we found that those trials that were registered had statistically lower risk of bias in the random sequence generation (88.2% vs. 66.1%, \( P = .03 \)), with no statistically significant differences in favor of trial registration in the rest of the quality assessment topics.

**DISCUSSION**

To our knowledge, this is the first study on methodologic quality of publications in the top fertility journals. More than 150 RCTs on infertility topics were published during the studied years, and most of them have low risk of bias; of significance, during the 6 reviewed years, we found that 6.8% of all the publications were RCTs. When analyzing only the three journals that published mainly clinical studies with primary data, the proportion of RCTs increased to 7.6%, but even in this case the proportion has increased significantly. In comparison, a 10-year review in aesthetic surgery journals only found 3.2% (11), and a similar study (12) found less than 2%.

In our study, only 10% of the RCTs reported to be funded by the pharmaceutical industry, whereas more than 25% had
either conflict of interest or unclear disclosure. Birkhahn et al. (13) have also studied the profile of author-reported funding of reports of original research published since 1994 in the four US peer-reviewed general emergency medicine journals. By 2003 there were 358 articles, with 36% funded (42% public, 30% foundation, 18% industry, and 10% multiple), representing more than triple of our findings. Conflicts of interest are usually reported in all the journals, although source of funding is not always systematically reported, even when the International Committee of Medical Journal Editors recommendation is to report it, not only when a conflict of interest is present but also if the supporting source had no such involvement (15). Because not all the studied journals require a funding source statement, it is likely that this item is under-reported. It is important that all journals call for funding source declaration, to allow the reader to decide whether this funding could bias the results.

It is encouraging that SRs were more commonly published in recent years, with a clear increment in some specific journals. However, the 1.5% of potential SRs found in our study should not be interpreted as if they were real systematic reviews with good-quality evidence. Additionally, when considering all the published studies analyzed, more than 90% of all the publications in the top five journals are neither SRs nor RCTs. However, it is important to mention that some journals are making an effort to improve the quality of their publications by including in their requisites for submission the appropriate checklist (i.e., CONSORT for RCTs) or a brief statement saying that the guidelines had been followed in the preparation of the manuscript. Given that trial registration was not widely promoted until 2004, it is not surprising that only 23% of the analyzed RCTs published in 2010 registered the trial previously (15). However, this figure should increase in the future to have better-quality research. It is remarkable that some of journals have also included in their “guide for authors” section the recommendation to do a clinical trial registration. In our study, trial registration was associated with a lower risk of bias, which confirms something that has been previously reported (16).

Bastian et al. (17) used a variety of data sources as MEDLINE and the Cochrane Central Register of Controlled Trials to estimate the numbers of trials and systematic reviews published from 1950 to 2007, showing that a plateau in growth has not yet been reached. Another study found a similar increasing trend for a specific journal (18). When looking at our data, with the exception of HR-Update’s trend in SRs publication, the journals included in our study did not follow this global increasing pattern of publication. However, our study design does not allow us to analyze whether this difference in this pattern has any specific explanation.

Our study was performed using a systematic review methodology, with two independent reviewers randomly allocated to each trial, using the strict Cochrane collaboration methods (7). This allowed us to perform an unbiased evaluation of the studies and to reach stronger conclusions. However, we recognize that our study has limitations. We only analyzed five major journals, and although they are the most often cited and have the highest impact factor, our conclusions only apply to these specific journals. Another significant criticism is that we used PubMed limits, which are not a precise tool to select SRs and RCTs. In 2011, Wieland et al. (19) showed that some of the RCTs are not tagged in MEDLINE; therefore, the figures presented in our study might be an underestimation of the real numbers. Wieland et al. showed that reporting and specific designs were not similar between tagged and untagged studies. One final limitation is that many published studies discussed topics that cannot be clearly evaluated with RCTs (e.g., diagnostic studies, risk factor studies). As a result, we cannot report the proportion of published RCTs out of all studies in which intervention questions could be potentially studied with an RCT design. If we had used as denominator only those studies whose question could have been answered with an RCT, the proportion of RCTs would have been higher.

Despite these limitations, our findings showing low risk of bias in most RCTs and an increasing number of SRs are certainly encouraging and could be relevant to clinicians, researchers, and editorial board members. Paying attention
to study designs, risks of bias, and conflicts of interest are important cornerstones in evidence-based medicine. Reporting results in an unbiased manner and obtaining quality data through well-known statements should always be encouraged, to move to a more complete, transparent, and adequate reporting and methodology, in particular for information related to treatment strategies. This initial effort to evaluate the research quality published in major reproductive medicine journals intends to set up a helpful baseline from which future improved protocols can be generated, to assess the quality of the reported studies in our subspecialty. In the process, this should benefit all those involved in the publication and acquisition of medical knowledge as we move forward with research in our field.

REFERENCES