

Reporting Standards and the Transparency of Trials

“Nobody sets out to write a bad paper,” said Curtis L. Meinert, PhD, professor of epidemiology and biostatistics at the Johns Hopkins School of Hygiene and Public Health in Baltimore. Yet far too many reports of clinical trials have lacked so much vital information that they are virtually unusable as guides for treating patients or planning policy. “There is a lack of detail in published reports that is due partly to authors and partly to editors,” said Meinert.

To bridge the gap between what a trial should report and what is actually published, a small group of researchers, epidemiologists, statisticians (including Meinert), and journal editors have met over the past 7 years to develop the Consolidated Standards of Reporting Trials (CONSORT), a set of guidelines that include a checklist and a model flow diagram, to help biomedical researchers report the results of randomized, controlled clinical trials fully and accurately. The first version of CONSORT appeared in 1996, and a major revision, which was completed in fall 2000, will soon be published. A detailed explanatory article giving the rationale for each item included in the reporting checklist and the empiric evidence that supports the importance of that item will accompany the revised CONSORT statement.

THE NEED FOR UNIFORMITY

“I felt for a long time that the medical literature was intimidating—a daunting pile of information which medical practitioners had a hard time using,” commented Edward J. Huth, MD, editor emeritus of *Annals*. According to Huth, MEDLINE, the familiar on-line bibliographic database of the U.S. National Library of Med-

icine (NLM), was designed more for researchers than for clinicians. “Practitioners want accurate, reliable, synoptic information—review articles by knowledgeable people,” he stated.

Authors preparing such review articles, however, must draw on the published results of many clinical trials. To avoid comparing apples and oranges, some uniformity in trial reporting is necessary. “Readers need details of studies to appraise them and know whether they can compare and pool data, know if they are equal and looking at the same things,” stated Ida Sim, MD, PhD, assistant professor of general internal medicine at University of California in San Francisco. “If you can’t compare trials, you can’t use them.” As a complement to, and outgrowth of, the CONSORT principle, Sim is in the process of developing an electronic databank called Trial Bank. This bank will contain trial information that captures and organizes study details in a consistent and easily accessible way.

To facilitate comparisons between trials, several groups have attempted to set standards for medical publishing. In 1978, a group of medical journal editors met in Vancouver, Canada, and issued formatting guidelines for articles submitted to their publications. This group, informally known as the Vancouver Group, eventually formed the International Council of Medical Journal Editors (ICMJE), which began work on a system of uniform reporting for medical journals that would require authors to meet only one set of editorial standards. Such standards would save time, too, because they would eliminate the need to send articles back to the author solely for formatting reasons, a particularly useful feature

when a manuscript is sent to a second journal after being rejected by the first. By 1997, more than 500 journals agreed to make the fifth edition of the ICMJE’s requirements their standard for manuscript submission. (Recently, these standards came full circle: At a meeting in Copenhagen in June 2000, the ICMJE voted to recommend CONSORT for use in reporting randomized clinical trials in its member journals.)

RCTs AND BIOMEDICAL RESEARCH

Randomized, controlled trials (RCTs) entered biomedical research barely half a century ago. The first, a study of the use of streptomycin in tuberculosis, was published by Bradford Hill in Great Britain in 1948 (*BMJ*. 1948;2:769-82). Why do RCTs hold such a special place in connecting biomedical research and clinical practice?

“Randomised controlled trials are the best way to compare the effectiveness of different interventions,” wrote Douglas G. Altman, DSc, another CONSORT member and professor of statistics in medicine at the Institute of Health Sciences in Oxford, United Kingdom. “It is thus entirely reasonable to require higher standards for papers reporting randomised trials than those describing other types of study . . . In essence the requirement is that authors should provide enough information for readers to know how the trial was performed so that they can judge whether the findings are likely to be reliable” (*BMJ*. 1996;313:570-1).

In theory, RCTs are designed to produce unbiased outcomes. However, if they are not done properly, biases may creep in and go undetected. In fact, RCTs seem to be

going through a midlife crisis of sorts (JAMA. 1998;279:319-20); many biases undermine their objectivity. According to Alejandro Jadad, MD, MPhil, of the University of Toronto, published RCTs tend to favor interventions for a number of reasons. For one, studies funded by industry-sponsored research have been shown to favor the sponsor and to be more poorly designed. For another, negative findings are published less often than positive ones. Finally, studies with positive results are more likely to be published in English. Jadad believes that this type of bias could be addressed by requiring registration of all clinical trials. Currently, registration is compulsory only for federally sponsored research, except in Spain, where registration of all trials is required.

Adding to the reader's problem is the incomplete presentation of important material, especially information pertaining to randomization procedures and their implementation. "Unfortunately, randomized controlled trials are anathema to the human spirit," wrote Kenneth Schulz, PhD, MBA, a biostatistician formerly with the Centers for Disease Control and Prevention and now vice president of quantitative sciences at Family Health International in Research Triangle Park, North Carolina. "Practitioners understand in theoretical terms the need for unbiased research, but once they are engaged in a trial they may find it too difficult to maintain a dispassionate stance. They may 'know' what treatment works better for patients and so want certain patients to benefit from that particular therapy; or they may want the result of a study to confirm what they

already believe" (CMAJ. 1995;153:783-6).

For this reason, said Schulz, trialists must clarify as much as possible the methods used to conceal randomization and patient allocation. Demanding that trial reports show how randomization and concealment were implemented is one way of bringing this crucial point to the attention of those conducting the trials. "In a randomized, controlled trial, it is not always possible to double blind," conceded David Moher, MSc, director, Thomas C. Chalmers Centre for Systematic Reviews, University of Ottawa, "but it is always possible to have adequate concealment and random generation. There's no methodological or practical reason why not."

ADVENT OF THE CONSORT GUIDELINES

In 1993 and 1994, two groups of researchers, biostatisticians, and editors began independently to explore the problem of lack of transparency in reporting RCTs and search for a solution. The groups—Standards of Reporting Trials (SORT) and the Asilomar Group—each developed a set of reporting standards. The SORT Group, Moher reported, began by looking for a scale for assessing the quality of clinical trials but found that so many reports were so poorly presented that it was impossible to tell whether the trials they described were of value. The group adopted an evidence-based approach for their standards and examined ways in which poor reporting of RCTs led to biased results. The Asilomar Group used a more experiential approach, placing more emphasis on statistical

aspects of reports rather than on issues of design and randomization.

Drummond Rennie, MD, deputy editor of *JAMA*, tested the early SORT standards by using as test cases two papers that had already been accepted for publication. Although the two rewritten papers met the SORT standards, the authors found them cumbersome to write in the full SORT format, and readers found them difficult to follow. Going back to the drawing board, the two groups met in Chicago and agreed on a combined statement—CONSORT—in 1995 (JAMA. 1996;276:637-9).

The CONSORT statement included two complementary ways of reporting information: a checklist and a flow diagram. The checklist consisted of 21 items that were felt to be necessary to inform the reader about what the researchers did during the trial and what they learned from it—their methods, results, and analysis. The flow diagram tracked patients through the various stages of the trial and detailed how many were eligible, were recruited, were randomly assigned to each arm, received treatment, completed the trial, or were lost to follow-up. However, not every aspect of a trial was necessarily included in the CONSORT reporting standards. Approval by institutional review boards and information about funding sources, for example, were not on the list.

Although the checklist addressed all aspects of a trial, it particularly emphasized the need for making randomization procedures explicit. The CONSORT statement explained the rationale:

"We recommend, for example, that RCTs should report how the allocation sequence was generated (e.g.,

computer generated) and concealed (e.g., in sequentially numbered, opaque, sealed envelopes) until the patient was randomized, something that is possible in all trials.” Schulz and colleagues have shown empirically that trials in which the allocation sequence had been inadequately concealed yielded larger estimates of treatment effects (odds ratios that differed, on average, by 30% to 40%) compared with trials in which the authors reported adequate allocation concealment (i.e., keeping the intervention assignments hidden from all persons involved with the trial until participants had been allocated to their respective study groups). According to the CONSORT statement, one possible interpretation is that some trials with inadequate reporting of allocation concealment actually had faulty randomization, and faulty randomization allowed the introduction of bias.

Whatever its intrinsic merits, CONSORT received powerful backing from the group that matters most to authors reporting scientific research—journal editors. The standards were adopted by *JAMA*, *Annals of Internal Medicine*, and the *British Medical Journal (BMJ)* and eventually by at least 70 other journals. In addition to the ICMJE (the Vancouver Group), such organizations as the Council of Science Editors and the World Association of Medical Editors have endorsed the CONSORT statement, ensuring that hundreds of journals now consider it their standard.

Not unexpectedly, CONSORT was met with cheers by those who saw the need for it but also with some grumbling. One letter to the *BMJ* complained of a North American bias among the originators, a dearth of

clinicians on the panel, a lack of proper respect for observational trials, and an “unduly prescriptive” approach (*BMJ*. 1997;314:1126). “Nobody likes to be told what to do,” commented Altman, a major participant in the current round of revisions, “but once the journals adopted CONSORT, it effectively became a de facto requirement for publication. However, nothing in CONSORT tells people how to run their trials, just how to report them. Transparency is the real issue.”

Physician-authors, Altman pointed out, are pressured by a system that requires them to publish research if they wish to advance their careers. According to Altman, CONSORT is aimed not at experts in controlled trials but at those who are looking for some guidance in preparing papers. “Authors value critical comments, even if it means rewriting the paper. The people who do clinical trials are not scientists; they’re primarily doctors. It is not a criticism to say that they are not trained in research. They don’t know how to include what the reader needs to know. Our primary intention is that whatever they did will be reported clearly to readers of the final paper. We’re not telling people to report any particular things in any particular order but that these are things you should take note of in your work,” he stated.

Even after 7 years and one major revision, CONSORT is viewed by its progenitors as a highly useful work in progress. Sylvan Green, MD, professor of epidemiology and biostatistics at Case Western Reserve University in Cleveland, Ohio, and CONSORT member, said he is an advocate of “large, high-quality, randomized clinical trials” and agrees with CON-

SORT’s premise. “To the extent that CONSORT is an outline of important principles and is constantly evolving and gets input from a wide variety of clinical trialists, then I think it is a good and useful idea. CONSORT is useful as a set of guidelines and suggestions, but not to force trial reports into the same framework. I would not want CONSORT to become too rigid,” he cautioned. Curtis Meinert is another believer who has some reservations. “Just putting a checklist in front of people is a valuable contribution,” he said. “It helps define and add crispness to clinical trials—which can use it.” Meinert’s chief objection has to do with analyses of subgroups, which can be useful but often contain such small numbers of patients that differences are too insignificant to report.

Another potential problem, according to Schulz, lies in the flow diagram because it takes up a large amount of space on the printed page. At the same time, a diagram can be a real plus. “Readers like it because it gives a graphic idea of the information, but surprisingly, editors like it too,” he commented. Even journals like *Lancet*, which publish shorter articles that describe studies with simpler designs, often use the flow diagram. Panel member Thomas Lang, MA, a consultant in scientific publications, said he wished that the statement had gone beyond statistics and transparency to consider more deeply the communications aspects of the checklist. “I’d ask about the audience, who would use it, and why,” he commented.

IMPACT OF THE GUIDELINES

What, exactly, has been the effect of the original CONSORT state-

ment? According to Altman, who often serves as a statistical referee for journal articles, two reports that study that question are in press in *JAMA*. He believes that authors have absorbed the ideas of the statement and that use of the checklist and flow diagram has led to more complete papers.

The recent CONSORT revision includes many small changes and friendlier language. An imperative tone to the checklist has been rendered more neutral. And as noted, a long explanatory document will also be available to flesh out the reasoning behind the statement's form and terminology, according to Moher. These documents will also be posted on the CONSORT Web site (www.consort-statement.org). Moher will invite users to submit good and bad examples of reporting of items on the checklist and will use the site as a laboratory and discussion area for future revisions. Moher has headed the revision effort from his base in Ottawa, relying on support from the Canadian Institutes of Health Research, some journals, the pharmaceutical industry, and the NLM.

The NLM's Ione Auston saw a confluence of interests between the Library and CONSORT. "The goals and objectives of the CONSORT project—to improve the quality of reporting of clinical research—relate to NLM's mission in supporting research in techniques and methods for recording, storing, retrieving, and communicating health information," she stated. "Journal editors reported that there were some problems in interpreting the original checklist items and also recommended additional

items to include in the checklist. If the current round of revisions respond as intended and effectively to these two issues, then they may be considered significant."

Other statements have been written that are similar to CONSORT. Epidemiologists have developed Meta-Analysis of Observational Studies in Epidemiology (MOOSE), and meta-analysts have developed Quality of Reporting of Meta-Analyses (QUORUM). The relationships among these groups are not accidental: Many experts, including David Moher, sat on all three panels.

In contrast to CONSORT, which is intended to guide the reporting of two-arm RCTs, MOOSE is designed to cope with data from "an existing database, a cross-sectional study, a case series, a case-control, a design with historical controls or a cohort design. Meta-analyses of observational studies present particular challenges because of inherent biases and differences in study designs; yet, they may provide tools for helping to understand and quantify sources of variability in results across studies" (*JAMA*. 2000;283:2008-12). The MOOSE checklist asks authors to report background and search strategy as well as methods, results, discussion, and conclusion.

The QUORUM statement has its own checklist and flow diagram, which are used to "report explicitly both the methods meta-analysts use to analyse the articles they reviewed, but also to report the methods used in the research articles they analysed" (*Lancet*. 1999;354:1896-900). In addition, QUORUM includes recommendations for a structured abstract and sections on validity assessment,

data abstraction, study characteristics, and quantitative data synthesis. The flow diagram depicts the RCTs that were included and excluded from the meta-analysis.

At the University of California, the Trial Bank being designed by Sim will take advantage of the Web's unlimited space. "Trial Bank not only reports results, it captures the design and execution of the study. It is structured, computer understandable, and not constrained in size," remarked Sim.

Trial Bank will ask authors of accepted trial manuscripts to place descriptions of their trials directly into an on-line repository (<http://rctbank.ucsf.edu>). Sim's CONSORT-Plus uses a controlled medical vocabulary (Unified Medical Language System) to make it more easily searchable. Eventually, she will evaluate Trial Bank with the help of systematic reviewers and practicing physicians. "So much money is devoted to clinical research," Sim stated, "but sloppy reporting systematically undermines our ability to use that research."

Although Altman believes that no available evidence shows that CONSORT has changed the way trials are designed, a firmer knowledge of what will be demanded in a report of a trial may eventually have an effect. "I hope that something filters upstream that could affect how they set up trials, but that is not the primary goal," he commented. The primary goal remains the clear presentation of randomized, controlled clinical trials in a way that makes their design, execution, and results evident to readers and useful for the practice of medicine.

—Aaron Levin