# Ophthalmic Randomized Controlled Trials Reports: The Statement of the Hypothesis

## CHUN FAN LEE, ANDY CHI ON CHENG, AND DANIEL YEE TAK FONG

• PURPOSE: To evaluate whether the ophthalmic randomized controlled trials (RCTs) were designed properly, their hypotheses stated clearly, and their conclusions drawn correctly.

• DESIGN: A systematic review of 206 ophthalmic RCTs.

• METHODS: The objective statement, methods, and results sections and the conclusions of RCTs published in 4 major general clinical ophthalmology journals from 2009 through 2011 were assessed. The clinical objective and specific hypothesis were the main outcome measures.

• RESULTS: The clinical objective of the trial was presented in 199 (96.6%) studies and the hypothesis was specified explicitly in 56 (27.2%) studies. One hundred ninety (92.2%) studies tested superiority. Among them, 17 (8.3%) studies comparing 2 or more active treatments concluded equal or similar effectiveness between the 2 arms after obtaining insignificant results. There were 5 noninferiority studies and 4 equivalence studies. How the treatments were compared was not mentioned in 1 of the noninferiority studies. Two of the equivalence studies did not specify the equivalence margin and used tests for detecting difference rather than confirming equivalence.

• CONCLUSIONS: The clinical objective commonly was stated, but the prospectively defined hypothesis tended to be understated in ophthalmic RCTs. Superiority was the most common type of comparison. Conclusions made in some of them with negative results were not consistent with the hypothesis, indicating that noninferiority or equivalence may be a more appropriate design. Flaws were common in the noninferiority and equivalence studies. Future ophthalmic researchers should choose the type of comparison carefully, specify the hypothesis clearly, and draw conclusions that are consistent with the hypothesis. (Am J Ophthalmol 2014;157: 254–259. © 2014 by Elsevier Inc. All rights reserved.)

ANY CLINICAL STUDY SHOULD HAVE A RESEARCH question. In randomized controlled trials (RCTs), the research questions usually are presented in the form of a statistical hypothesis and are answered by comparing the outcome measures between different treatment groups. From a statistical point of view, the research question is said to be answered only when the null hypothesis is rejected. If the null hypothesis cannot be rejected, one can say only that it is inconclusive, rather than claiming that the null hypothesis is accepted or confirmed. For example, in a superiority trial that aims to test if one treatment is more effective than another, a significant result reflected by a small enough $P$ value indicates superiority, but an insignificant result does not imply that the 2 treatments are similar, because such a result may occur as a result of a small sample size.

A prospectively defined hypothesis is one of the items required by the Consolidated Standards of Reporting Trials statement in reporting clinical trials.[1,2] Despite its importance, the hypothesis often is described inadequately. Scherer and Crawley reviewed ophthalmic RCTs published between 1991 and 1994, and they found that less than 50% of the reviewed articles in *Archives of Ophthalmology*, approximately 25% in *Ophthalmology*, and slightly more than 50% in the *American Journal of Ophthalmology* stated the study hypothesis.[3] Sánchez-Thorin and associates reported that only 1 of 24 RCTs published in *Ophthalmology* in 1999 explicitly described the hypothesis.[4] These ophthalmic RCTs were reported at least 10 years ago, and there has been no update since the revision of CONSORT statement in 2001.[1]

In terms of how the treatment groups are compared, statistical comparisons can be classified into 3 types of design: superiority, noninferiority, and equivalence. Superiority is tested when one aims to show a study treatment has better (1-sided) or different (2-sided) performance than another treatment. A noninferiority study tests whether the study treatment is at least as good as the control treatment with respect to some particular aspect such as efficacy.[5] The study treatment usually offers benefits in other aspects over the control treatment such as increased safety, better quality of life, higher compliance, less toxicity, or reduced cost.[6] The control treatment in a noninferiority study must be an active treatment.[7] An equivalence study aims to show one treatment performing equally effectively as another.[5] These 3 types of statistical comparisons require different analysis techniques.[8] Moreover, the sample size required in a superiority study often is different from that

of a noninferiority or an equivalence study. Therefore, the decision of whether a superiority, noninferiority, or equivalence design is to be used should be made in the designing phase of a study. The objective of this review was to examine whether recently published ophthalmic RCTs were designed properly, whether their hypotheses were stated clearly, and whether the conclusions were drawn consistent with the hypotheses.

## METHODS

• **SELECTION OF RANDOMIZED CONTROLLED TRIALS:** We hand-searched for RCTs published between January 2009 and December 2011 in the 4 top-ranking general clinical ophthalmology journals according to their impact factor in 2009: the *American Journal of Ophthalmology*, *Archives of Ophthalmology*, the *British Journal of Ophthalmology*, and *Ophthalmology*. Abstracts of articles with the word *random* (or any other form) in the title or abstract were read independently by 2 authors (C.F.L. and A.C.O.C.). Those that were considered to be nonrandomized studies, cohort studies, cross-sectional studies, case-control studies, case series, systematic reviews, meta-analyses, and animal studies were excluded. For the remaining articles, the methods sections were examined carefully. Articles reporting more than 1 RCT and those reporting only results of 1 treatment group, study design, baseline data, or a combination thereof also were excluded.

• **DATA EXTRACTION:** The 2 authors independently extracted data of the eligible studies using a standardized Excel (Microsoft, Redmond, Washington, USA) template by reading through the entire article. These data included:

1. The clinical objective of the study. The CONSORT statement recommends that authors state the research questions that were designed to be answered, that is, the clinical objective.[1,2] The 4 included journals require a structured abstract in which the purpose, aim, or objective of the reported trial is stated. If no objective statement could be found directly in the main text, the objective stated in the abstract was assessed instead.
2. The primary hypothesis and its type. We defined in this review a hypothesis as a specific statement that allows one to construct the statistical evaluation. Based on the evaluation, each trial can be classified by the way of comparison into a superiority, noninferiority, or equivalence study. Therefore, studies that did not describe how the treatments were compared were regarded as having no hypothesis specified. If no hypothesis was described in the objective statement, we determined the type of hypothesis according to the descriptions in the methods and results sections. For example, if the methods section stated that a $t$ test was used or the results section

presented the $P$ value for a $t$ test, it was determined that the comparison was a superiority test.
3. The number of treatment groups, the treatment used in each group, and the descriptions of the treatments. The trials were classified by what treatment groups were compared: whether the study treatment was compared with no intervention, placebo, or another active treatment or whether the study treatment was used as an adjunctive treatment on top of some base treatments.
4. Whether the trial was sponsored by industry. The sections of financial disclosure, financial support, funding, competing interests, and acknowledgement of each reported study were reviewed. The studies then were classified as industry sponsored, not industry sponsored (sponsored by government or nonindustry organizations for academic purpose), or having no sponsor according to the description in these sections.
5. The interpretation of the results and the conclusions drawn by the study authors.

• **STATISTICAL ANALYSIS:** The numbers of studies with different types of comparisons and treatment groups were counted and the percentages were calculated. Because we did not intend to compare the 4 journals, no statistical comparison between the journals was performed.

## RESULTS

A TOTAL OF 559 ARTICLES CONTAINING THE WORD *random* in the title or abstract were identified from the 4 journals (Table 1). Three hundred fifty-three articles were excluded after detailed analysis of the methods section. These were mainly articles with study designs other than an RCT, for example, a survey that selected subjects with random sampling. Three other studies that were not exactly RCTs, including 2 that performed randomization by alternation and 1 that allocated the left eye to one group and right eye to another group, also were included. Finally, 206 studies were included in our assessment.

One hundred ninety-one of the eligible studies (96.6%) stated the clinical objective. One hundred seventy-one studies (83.0%) stated the objective both in the abstract and in the main text, whereas 22 studies (10.7%) stated the objective only in the abstract, and 6 studies (2.9%) stated the objective only in the main text. The remaining 7 studies (3.4%) did not mention the clinical objective, but aimed to "report the outcomes" or "describe the results" of the trial.

Table 2 summarizes the number of studies with and without the hypothesis clearly stated. Fifteen studies (7.3%) explicitly stated the hypothesis (eg, "we hypothesize," "test this hypothesis that," etc), and another 41 studies (19.9%) did not use the term *hypothesize* (or in another form), but instead used some phrases that implied the type