JACC: CARDIOVASCULAR INTERVENTIONS
VOL. ■, NO. ■, 2017
© 2017 BY THE AMERICAN COLLEGE OF CARDIOLOGY FOUNDATION
ISSN 1936-8798/$36.00
PUBLISHED BY ELSEVIER
http://dx.doi.org/10.1016/j.jcin.2016.11.041

EDITORIAL COMMENT

What Do Noninferiority Trials Say About Coronary Stents?*

John A. Bittl, MD

When investigators are trying to prove that a new treatment is superior to an existing one, they commonly use statistical methods that prevent them from saying what they want to say (1). Before showing that a new treatment is better than an old one, investigators must first fabricate a null hypothesis that says there is no difference between the new and old treatments. After a clinical trial has been completed, a p value is calculated from a mathematical simulation that represents how frequently the observed data would have occurred if the trial could have been repeated a large number of times and the null hypothesis were true. In the frequentist framework, the p value represents the proportion of trials that would have produced similar results—based on a large series of repetitions that does not occur, conditional on a null hypothesis that no one truly believes (2). A p value of less than an arbitrary cutoff of 0.05 allows investigators to conclude that, "Either an exceptionally rare chance has occurred or the [null hypothesis] is not true," (3) although in fact there may be no relation between the probability of the null hypothesis and the plausibility of an alternative hypothesis. Generating a point estimate and 95% confidence interval helps to define the treatment effect but does little to reverse the inverted logic of frequentist statistics (4,5).†

When investigators are trying to prove that a new treatment is noninferior to a preexisting therapy, the statistical methods become more tortuous and nonintuitive (5). To compare 2 treatments using a marginal analysis (6), investigators postulate a "null" that describes a difference between the 2 treatments, as defined by a noninferiority margin -Δ (5). The goal of a noninferiority trial is to compare the 2 treatments, obtain a 95% confidence interval with a lower bound that does not cross the noninferiority margin, and satisfy the alternative hypothesis that the difference between the new and old treatments is less than -Δ. In the backward world of noninferiority, the null hypothesis contains a treatment difference, whereas the alternative hypothesis seems "null" (5).

Despite the convoluted statistical design, noninferiority trials have been popular for comparing stents. In this issue of *JACC: Cardiovascular Interventions,* 2 reports (7,8) describe several such studies. Yeh et al. (7) summarize the experience with a durable-polymer zotarolimus-eluting stent (DP-ZES) in a series of noninferiority trials and registries known as the Resolute program. The investigators report that the 5-year cumulative incidence of definite or probable stent thrombosis was 1.2%, which consisted of a rate of 0.7% at 1 year and an annualized rate of 0.1% thereafter. Although the Resolute program (7) may have enrolled patients with higher risk than, say, the DAPT (Dual Antiplatelet Therapy) trial (9), a study that used a superiority design, with more patients in the Resolute program having high-risk baseline characteristics such as prior myocardial infarction (29% vs. 22%) or diabetes (37% vs. 31%) than patients in the DAPT trial (9), the annualized rate of very late stent thrombosis was 3 to 10 times lower in the Resolute program than in the DAPT trial. Does this discrepancy reflect differences in stent safety, trial design, or ascertainment of events?

*Editorials published in *JACC: Cardiovascular Interventions* reflect the views of the authors and do not necessarily represent the views of *JACC: Cardiovascular Interventions* or the American College of Cardiology.

From the Interventional Cardiology Group, Munroe Regional Medical Center, Ocala, Florida. The author has reported that he has no relationships relevant to the contents of this paper to disclose.

†The 95% confidence interval, like a p value, is commonly misunderstood as a direct probability statement, but there is not a 95% guarantee that the true treatment effect lies within the confidence interval. Instead, the traditional 95% confidence interval describes the hypothetical performance of a large number of similar trials, and like the theoretical exercise for defining a p value, the process would generate a large number of confidence intervals, of which 95% will contain the true odd ratio but the actual probability that a particular confidence interval contains the true odd ratio cannot be stated (5).

Investigators involved in superiority trials are rewarded for achieving high enrollments and capturing all events to maximize the chance that they will find a difference. Not meant to be a criticism, but investigators involved in noninferiority trials or registries are not penalized for losses to follow-up or incomplete ascertainment of events, because such deficiencies could increase the chance of finding no difference (5). Along these lines, it was reassuring that the Resolute program (7) included studies that reported 2-year follow-up adherence as high as 98.1% and observed 5-year rates of 5.0% for cardiac death and 4.4% for target vessel myocardial infarction that were more in line with the 18-month rates of 0.9% for cardiac death and 3.0% for myocardial infarction in the DAPT trial (9).

In a second paper In this issue of *JACC: Cardiovascular Interventions*, Raungaard et al. (8) report the 3-year results of the SORT OUT (Scandinavian Organization for Randomized Trials with Clinical Outcome) VI trial, a noninferiority trial that compared the DP-ZES with a biodegradable-polymer biolimus-eluting stent (BP-BES) in 2,999 patients. The original publication (10) reassuringly reported follow-up in 99.7% of enrolled patients but found event rates for the primary endpoint of 5.3% in the DP-ZES group and 5.0% in the BP-BES group that were both lower than the expected rate of 6.5% and differed from each other by less than the prespecified absolute risk margin of 2.5% (10). To affirm noninferiority of the DP-ZES compared with the BP-BES, the investigators (8) now report no difference in major adverse cardiac events (8.6% vs. 9.6%), death (2.7% vs. 3.4%), or very late stent thrombosis (0.4% vs. 0.7%) at 3 years, respectively.

Noninferiority trials have become an integral part of the interventional cardiology evidence base and do not necessarily prove that industry has hacked frequentist statistics to promote a raft of "me-too" coronary stents. Although most investigators would prefer to declare that a new stent is better than an old stent, some practitioners might be willing to sacrifice a little efficacy in exchange for improved deliverability, lower cost, or shorter courses of DAPT, although the present studies (7,8) do not show directly whether the DP-ZES or the BP-BES offers such advantages.

A drawback common to both noninferiority and superiority trials involves changes in the standard of care and selection of the wrong comparator. In noninferiority trials this may lead to the problem of "techno-creep" (5), in which progressively inferior devices are shown to be "not unacceptably worse" (5) in sequential comparisons (11). In this regard, it is important to recognize the emerging evidence (12,13) that a different coronary stent, a durable-polymer everolimus-eluting stent containing an antithrombotic fluoropolymer (14) on a thromboresistant platinum chromium metallic backbone (15), is the current standard.

The present studies (7,8), along with a separate noninferiority trial (16), provide a level of reassurance that the DP-ZES is not unacceptably worse than multiple biodegradable-polymer stents. But, given the ever-changing evidence base in interventional cardiology and the convoluted statistics of noninferiority trials, some practitioners may wonder how a DP-ZES would stand up against a durable-polymer everolimus-eluting stent or question whether better methods of comparison might be found. Future efforts could explore the possibility of using Bayesian approaches, which predated frequentist approaches by at least 150 years (1,4) and are firmly rooted in the science of probability (17). A central component of the Bayesian approach is the Bayes factor, which in its simplest form is a likelihood ratio, which has been applied widely in diagnostic testing to calculate the predictive value of a positive result and in its logarithmic form $\log_e$(OR) labeled by Alan Turing as the "weight of evidence" to crack the Nazi Enigma codes in the Second World War (4). As compared with p values, which greatly overstate the evidence against the null hypothesis (18), Bayes factors range from 0 to ∞ and constitute a potentially superior approach that weighs the relative likelihood of 2 different hypotheses, with small values close to 0 simultaneously providing strong evidence against a true null hypothesis *and* for an alternative hypothesis (1,4,18). A fresh look at an old approach might allow investigators to make direct probability statements from their data and say precisely what they mean to say.

REPRINT REQUESTS AND CORRESPONDENCE: Dr. John A. Bittl, MD, Interventional Cardiology Group, Munroe Regional Medical Center, 1221 SE 5th Street, Ocala, Florida 34471. E-mail: jabittl@mac.com.

REFERENCES

1. Jeffreys H. Theory of probability. 3rd ed. Oxford: Clarendon Press, 1961.

2. Goodman SN. Toward evidence-based medical statistics. 1: the P value fallacy. Ann Intern Med 1999;130:995–1004.

3. Diamond GA, Kaul S. Prior convictions: Bayesian approaches to the analysis and interpretation of clinical megatrials. J Am Coll Cardiol 2004;43:1929–39.

4. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian Approaches to Clinical Trials and Health Care Evaluations. Chichester, UK: Wiley, 2004.

5. Schumi J, Wittes JT. Through the looking glass: understanding noninferiority. Trials 2011; 12:106.

6. Kaul S, Diamond GA, Weintraub WS. Trials and tribulations of non-inferiority: the ximelagatran experience. J Am Coll Cardiol 2005;46:1986–95.

7. Yeh RW, Silber S, Chen L, et al. Five-year safety and efficacy of Resolute zotarolimus-eluting stent: the RESOLUTE global clinical trial program. J Am Coll Cardiol Intv 2017;11:00.

8. Raungaard B, Christiansen EH, Bøtker HE, et al. Comparison of durable-polymer zotarolimus-eluting and biodegradeable-polymer biolimus-eluting coronary stents in patients with coronary artery disease: 3-year outcomes in the randomized SORT OUT VI Trial. J Am Coll Cardiol Intv 2017;11:00.

9. Mauri L, Kereiakes DJ, Yeh RW, et al. Twelve or 30 months of dual-antiplatelet therapy after drug-eluting stents. N Engl J Med 2014;371:2155–66.

10. Raungaard B, Jensen LO, Tilsted H-H, et al. Zotarolimus-eluting durable-polymer-coated stent versus a biolimus-eluting biodegradable-polymer-coated stent in unselected patients undergoing percutaneous coronary intervention (SORT OUT VI): a randomised non-inferiority trial. Lancet 2015; 385:1527–35.

11. Kaul S, Diamond GA. Good enough: a primer on the analysis and interpretation of noninferiority trials. Ann Intern Med 2006;145:62–9.

12. Palmerini T, Biondi-Zoccai G, Della Riva D, et al. Clinical outcomes with bioabsorbable polymer-based versus durable polymer-based drug-eluting stents and bare metal stents: evidence from a comprehensive network meta-analysis. J Am Coll Cardiol 2014;63:299–307.

13. Kang SH, Chae IH, Park JJ, et al. Stent thrombosis with drug-eluting stents and bioresorbable scaffolds: evidence from a network meta-analysis of 147 trials. J Am Coll Cardiol Intv 2016;9:1203–12.

14. Kolandaivelu K, Swaminathan R, Gibson WJ, et al. Stent thrombogenicity early in high-risk interventional settings is driven by stent design and deployment and protected by polymer-drug coatings. Circulation 2011;123: 1400–9.

15. Eppihimer MJ, Sushkova N, Grimsby JL, et al. Impact of stent surface on thrombogenecity and vascular healing: a comparative analysis of metallic and polymeric surfaces. Circ Cardiovasc Interv 2013;6:370–7.

16. von Birgelen C, Kok MM, van der Heijden LC, et al. Very thin strut biodegradable polymer everolimus-eluting and sirolimus-eluting stents versus durable polymer zotarolimus-eluting stents in allcomers with coronary artery disease (BIO-RESORT): a three-arm, randomised, non-inferiority trial. Lancet 2016;388:2607–17.

17. Bertsekas DP, Tsitsiklas JN. Introduction to Probability. 2nd edition. Belmont, MA: Athena Scientific, 2008.

18. Goodman SN. Toward evidence-based medical statistics. 2: The Bayes factor. Ann Intern Med 1999;130:1005–13.

KEY WORDS Bayesian analysis, clinical trials, percutaneous coronary intervention, randomized controlled trials, p value, stent polymers