Any testing of medical treatments is an exercise in comparison. In a typical clinical trial, 2 treatments are compared to determine which is better or if both are the same. Trials designed to address whether one treatment is better than the other may be called superiority trials, whereas those designed to show that 2 treatments are the same are called equivalence trials. However, the design of both types of trials should depend on the uncertainty principle—a fundamental ethical and scientific principle for conducting randomized controlled trials.1

The article by Staszewski et al2 in this issue of THE JOURNAL is a randomized controlled equivalence trial that compares a triple nucleoside regimen of abacavir-lamivudine-zidovudine with a more conventional regimen of indinavir-lamivudine-zidovudine in treatment-naive patients infected with human immunodeficiency virus (HIV). Although the authors conclude that these 2 regimens are equivalent in achieving the primary end point of reducing plasma HIV RNA levels to below 400 copies/mL, several factors make the interpretation of this study and other equivalence trials particularly difficult.

In planning a clinical trial of a new intervention, 2 main issues must be addressed. The first is the fundamental ethical question of whether the use of the new intervention is justified. The second is the choice of the appropriate control group.3 Both issues are fundamentally related to the pre-existing knowledge about the therapeutic value of the treatments to be compared. This is an important reason that clinical trials should be preceded by a systematic review to assess the status of this knowledge, and should be reported with a discussion of an updated review including the trial’s results.4 The trial would not be justified if one of the treatments to be assessed is known to be superior to the other. A clinical trial is only justified if the patient and clinician are not certain about which treatment to choose from the available options. If they are uncertain (indifferent) about the relative value of the treatments, it is time for a trial.5 This is not only because the trial will help resolve this uncertainty but also because it is the fairest way to choose the treatment for the patient. Patients enrolled in the study have a 50% chance of receiving the better treatment and the overwhelming weight of evidence is that they will fare better while participating in the trial (regardless of the treatment they are allocated to) than while outside of it.6 This realization forms a basis for the scientific and ethical underpinnings for the design and conduct of randomized trials, expressed in the term uncertainty principle, which states that a patient should be enrolled in a randomized controlled trial only if there is substantial uncertainty about which of the trial treatments would benefit the patient more.7

Basing trials on the uncertainty principle also addresses another important issue in the design of a clinical trial—the choice of an adequate comparator for the intervention under investigation.8 Studies in which the intervention and the control or comparison group are known in advance to be nonequivalent in their effects on the main outcomes of interest violate the uncertainty principle.9 Even if a study is properly reported,9 extra caution might be needed in its interpretation if the choice of the comparison treatment was not based on uncertainty about the relative value of the treatments being assessed.8,10

Uncertainty can have many grades ranging from simply not knowing11 to maximum uncertainty (also known as equipoise)11,12 about the relative benefits and harms of the treatment alternatives. The uncertainty might be in the mind of the patient, the clinician, or the community.12 Most clinical trials are assessments of superiority and start with the statement of a null hypothesis of no difference between 2 therapies. That is, prior research should not have proved a difference between the alternative treatments in the outcomes to be assessed. The trial is designed to reject the null hypothesis by showing that there is a difference between the treatments. Since the null hypothesis can never be proven, but only rejected,10 alternative hypotheses (ie, that one treatment is better) are not assessed directly, but are accepted if the probability that the observed results are consistent with a different hypothesis (eg, that 1 treatment is better) is above a predetermined level.11

Author Affiliations: Interdisciplinary Oncology Program, Division of Blood and Bone Marrow Transplantation, H Lee Moffitt Cancer Center and Research Institute, University of South Florida, Tampa (Dr Djulbegovic); and Cochrane Centre, NHS Research and Development Program, Summertown Pavilion, Oxford, England (Dr Clarke).

Corresponding Author and Reprints: Benjamin Djulbegovic, MD, PhD, H Lee Moffitt Cancer Center and Research Institute, University of South Florida, 12902 Magnolia Dr, Tampa, FL 33612 (e-mail: djulbebm@moffitt.usf.edu).

©2001 American Medical Association. All rights reserved.
Several factors may make the interpretation of equivalence trial results particularly difficult. When designing a superiority trial, a power calculation and sample size determination are performed to assess the probability that a given difference is obtained by chance. In equivalence trials, this difference ideally would be zero, although a proof of exact equality is not possible. In practice, this issue is resolved by defining an arbitrary practical equivalence margin, also called the non-inferiority margin. To detect this difference, on average, equivalence trials usually will require a 10% larger sample size in comparison with conventional superiority trials. The null hypothesis would be rejected if the upper limit of the confidence interval for the difference between the treatments is smaller than this predefined margin. Setting of the margin is critical and should be chosen on the basis of excluding a clinically important difference between the treatments. However, the definition of what constitutes such a difference may vary widely for each patient and clinician and might fall below the margins set by the trialist.

In this study, Staszewski et al set the limit for the difference at 12% for their primary end point (a plasma HIV RNA level of 400 copies/mL at week 48) based on discussions among clinical investigators of the study and with officials from the Food and Drug Administration.

Once the study results are obtained, a key question for the interpretation of equivalence trials revolves around whether both treatments were effective, or whether the result indicates that both treatments were ineffective. This could also be a feature in the interpretation of superiority trials and is one reason that a placebo or no treatment control arm would be used (if an active control treatment does not exist). Of course, a substantial amount of historical data indicate that the lack of a treatment arm would not be appropriate for studies such as the trial by the Staszewski et al involving patients with HIV infection who require treatment. Nonetheless, its interpretation remains problematic. Absence of evidence (of a difference) must not be confused with evidence of absence (of a difference). The observation of a lack of a difference between 2 treatments cannot automatically be used as evidence of equivalence.

An additional problem is that common techniques used to minimize bias in clinical trials are less useful in equivalence trials. In conventional trials, use of randomization, blinding, and intent-to-treat analysis serve 1 purpose: to ensure comparability between the 2 groups in all respects other than the study treatment so that any outcomes that differ between the groups could only be due to the study treatment or to chance. However, when the intent is to show that the study treatment is identical to control, techniques that ensure similarities between 2 groups are less helpful. Indeed, in the study by Staszewski et al, results of the intent-to-treat analysis differ from the as-treated analysis, with the more conventional regimen of indinavir-lamivudine-zidovudine actually appearing to do better. Although reasons for this difficulty are difficult to discern without examination of the actual data in the trial, a power calculation and sample size determination are performed to assess the probability that a given difference is obtained by chance. In equivalence trials, this difference ideally would be zero, although a proof of exact equality is not possible. In practice, this issue is resolved by defining an arbitrary practical equivalence margin, also called the non-inferiority margin. To detect this difference, on average, equivalence trials usually will require a 10% larger sample size in comparison with conventional superiority trials. The null hypothesis would be rejected if the upper limit of the confidence interval for the difference between the treatments is smaller than this predefined margin. Setting of the margin is critical and should be chosen on the basis of excluding a clinically important difference between the treatments. However, the definition of what constitutes such a difference may vary widely for each patient and clinician and might fall below the margins set by the trialist.
trial, there were large differences in the proportions of patients for the intent-to-treat and as-treated analyses in the 2 treatment groups. For example, for the primary end point, the difference in proportions of patients analyzed was 55% (133/262 vs 125/145 in the abacavir-lamivudine-zidovudine group) and 52% (136/265 vs 130/139 in the indinavir-lamivudine-zidovudine group). Therefore, the apparent equivalence reflected in the intent-to-treat analysis might simply be due to a dilutional effect of comparing 2 groups of patients whose actual treatments did not differ much.

The design of a clinical trial should be a function of the uncertainty principle, which should underpin both superiority trials and equivalence trials. However, the former are usually designed in the hope that one treatment will prove better than the other, whereas the latter are designed in the hope that both are the same. The main impetus for an equivalence trial is the notion that proving equal efficacy may enable patients to have treatments that are not more effective than existing ones, but are better for some other reason. However, if 2 treatments are shown to have equal efficacy, the evidence that one is better should also be of the highest standard possible. Clinicians should be cautious about arguing for equivalence based on a randomized trial of efficacy, and then using arguments about toxicity (or some other end point that was not formally assessed in the trial) as a basis for suggesting the superiority of one of the treatments. Problems identified with the interpretation of equivalence trials should not necessarily argue against their conduct. Rather, such trials need to be designed and reported in a transparent and explicit fashion, to acknowledge that they are not really equivalent to superiority trials.

REFERENCES


6. How do the outcomes of patients treated within randomised control trials compare with those of similar patients treated outside these trials? Available at: http://hiru.mcmaster.ca/ebm/trout/. Accessibility verified February 1, 2001.


10. Djulbegovic B, Bennett C, Lyman G. Violation of the uncertainty principle in conduct of randomized controlled trials (RCTs) of erythropoietin (EPO). Blood. 1999;94:399A.


