Equipoise abandoned? Randomization and clinical trials

Randomized, clinical trials are the strongest method we have to define the efficacy of new drugs. Recently, however, novel agents that show remarkable responses with minimal toxicity in patients with advanced cancer have been discovered [1], and these present a challenge for the ethical conduct of randomized trials. Consequently, there is a debate on the randomized trial as the gold standard for bringing these agents to approval.

For several decades, the standard for determining that a randomized trial is ethical has been a matter of discussion. Some doctors espouse the uncertainty principle, whereby randomization to treatment is acceptable when an individual doctor is genuinely unsure which treatment is best for a patient. Others believe that ‘clinical equipoise,’ reflecting collective professional uncertainty about treatment, is the soundest ethical criterion [2, 3]. Both sides agree that uncertainty as to the more effective therapy is necessary for a randomized trial to be ethical, the point of contention being only whether it is individual physician uncertainty or collective uncertainty of the medical community. Until the recent proposal to reject the equipoise (balance of uncertainty) standard [4], it has been widely accepted that doubt as to the better arm must exist at the start of the trial, and safeguards are also required so that there is intermittent monitoring of the trial by an independent committee, who can stop the trial should one of the arms prove to be superior.

The problem arises when substantial evidence suggests that a new treatment has strikingly superior efficacy, and the equipoise standard may therefore not be met. To address this concern, it has recently been proposed that, to inform coverage decisions for populations, payors may require rigorous assessment of a new agent beyond the point at which physicians and informed patients would choose it over customary care [4]. This argument is provided to justify randomized trials with arms that are anticipated to have very different outcomes [4]. The suggestion that equipoise should be disregarded in this context was based on several considerations [4]:

(i) imprecision in defining the concept of equipoise,
(ii) the susceptibility of expert opinion to bias
(iii) limitations of determining efficacy on the basis of surrogate outcomes used in non-randomized trials,
(iv) high costs of new treatments,
(v) tendency toward premature termination of randomized clinical trials as a result of the equipoise mandate.

In this article, we discuss these considerations, and dispute the contention that equipoise should be abandoned as the arbiter of the merit/legitimacy of a randomized trial.

The first issue that Miller and Joffe [4] point out in their argument for abandoning equipoise is that there is imprecision in defining the concept of equipoise, i.e. there is no standard way to elucidate when a consensus exists in the medical community for the superiority of a certain treatment. Therefore, the need of government agencies and insurance payors for more rigorous population-based data justifies a randomized trial. We contend that the argument that the lack of consensus amongst medical experts regarding the need for a randomized trial can be used as a rationale to abandon equipoise is flawed; this argument is faulty because it assumes that there would be better consensus amongst government experts and insurers regarding the need for a randomized trial than among medical experts, and there is no evidence to support such agreement amongst the former groups. Furthermore, it neglects the obvious conflict of interest that payors have in regard to making decisions as to when they should be obligated to cover a treatment.

According to Miller and Joffe, the susceptibility of expert opinion to bias is a second issue that confounds equipoise [4].
Indeed, it has been asserted that the clinical equipoise standard relies on ‘mere’ expert opinion to determine the permissibility of randomized trials [4]. Joffe and Miller [4] rightly point out the fallibility of expert opinion, and ubiquitously touted examples of treatments felt by experts to be beneficial that later proved harmful. However, because medical experts are sometimes fallible, it should not be inferred that they are always or even usually wrong. Indeed, there are numerous studies demonstrating that patients treated by experienced physicians with a high degree of proficiency are likely to do better than are patients who are cared for by physicians with less expertise [5]. Finally, as mentioned above, presumably government and insurer payor experts would not be immune to fallibility or bias; indeed, their biases may be severe considering their conflicts of interest, since they are obligated to pay for the therapies once approved, and so may seek to delay their approval.

The third confounder for non-randomized trials relates to the limitations of determining the efficacy on the basis of common surrogate end points—response and time to progression—instead of the ‘gold’ standard of survival. It is well known that responses in many phase II trials do not predict for a successful phase III study. However, there is also extensive literature, suggesting that response and progression-free survival can be useful surrogates, especially when utilized correctly and with threshold values [6]. Johnson et al. [6] described a threshold level above which improvement in either response rate (between 18% and 38% [increase in the response rate]) or progression-free survival (incremental gain needed was between 1.8 and 3.3 months) predicted for improvement in overall survival. Furthermore, a survey of all 31 anticancer drugs approved in the United States over a period of 34 years based on the response rate or progression-free survival, without a randomized trial that utilized a comparator arm, demonstrated that these agents fared well, with all but one showing long-term safety and efficacy [7]. It should also be kept in mind that the end points of randomized trials are not without problems. For instance, survival may be confounded by subsequent therapies, and this end point requires large studies and long periods of follow-up.

The high costs of new treatments have been promoted as a fourth justification for the need for a randomized trial that may not meet the equipoise benchmark [4]. It has been further claimed that, when cost is considered, randomized trials may be deemed to be ethical in order to generate the rigorous knowledge needed to guide health policy decisions despite lack of equipoise, so that new treatments with an unfavorable risk–benefit ratio or marginal benefits are not approved [4]. However, it would be reasonable to maintain that cost considerations should not play any role in determining equipoise, nor consequently in assessing the ethics of randomized trials. Furthermore, even if cost considerations are permitted as an arbiter of ethics, the assertions regarding costs fail to address the ultimate higher price of treatment because of the extra 20–100 million dollars and, perhaps more importantly, 2–5 years spent on the randomized trial(s) [8]. In addition, the cost in lives and lost productivity during this time period, when a drug with high response rates is not available because it is undergoing a randomized trial, should be part of the calculation, and may drive up, rather than down, health care expenditures.

Even more importantly, the model of the randomized trial is one that has been especially conducive to approval of numerous expensive anticancer drugs, which show improved survival that more often than not ranges from less than a few weeks to 3 months [9–11]. It is also plausible that randomized, controlled trials have led to the approval of ineffective therapies. Indeed, a relatively common phenomenon in drug development is carrying out randomized trials in cancers where there is a large and potentially lucrative market, even though there may have been few to no responses in those cancers during phase I to II development. If one runs multiple comparisons, and sets the P value at 0.05 for significance, there can be a ‘multiplicity issue,’ that is the phenomenon of finding that some of the comparisons (5% to be precise) show significant differences simply because so many comparisons were carried out. Because of the large numbers of randomized trials that are carried out, especially in patients with common cancers, it appears conceivable that some drugs without true efficacy have been approved. Biologically, it might be rational to surmise that drugs that do not result in tumor regression, but show improved survival for a few weeks in a large, randomized trial, are the ones that are most liable to being approved on the basis of a false-positive randomized trial.

A fifth issue that is levied against the equipoise standard is the fact that maintaining equipoise promotes premature termination of randomized, clinical trials [4]. In particular, when extremely positive results are seen and a boundary is crossed, the data monitoring committee for a randomized study may recommend stopping the trial and releasing the results. Although some authors contend that early stopping is associated with inflation of the treatment effect, others state that, with a well-designed interim-monitoring plan, stopping early has a negligible impact on estimation of the true level of benefit [12]. Further, the contention that a trial should be continued beyond a certain boundary that is believed to indicate superiority for one of the arms, by necessity means withholding that information from patients. Such a practice appears to be unethical, and might well raise health care costs because of the legal liabilities incurred.

The issue of informed consent process must also be considered. It has been proposed that it is ethical to request patient participation in a randomized trial, even when the majority of the medical community is convinced of the value of the new drug; the informed consent mandate in this instance would be covered by letting the patient know that they are being asked to enroll in order to gain new knowledge that would help future patients [4]. However, in our experience, research participants rely on their discussion with the physician–investigator to make their decisions. If the medical community including the physician–investigator believes that the doubts remaining about the comparator arms are few, the physician–investigator should be obligated to share that information with the patient. Indeed, regardless of who obtains consent, patients should be fully informed regarding consensus medical opinions.

The limitations of randomized, clinical trials must also be taken into account. While there is little doubt that randomized, controlled trials have significant and well-known advantages, they also have drawbacks. These include the difficulty in generalizing the results of research done in such well-controlled populations. In addition, Booth et al. [13], in a comprehensive

2472 | Kurzrock and Stewart
Volume 24 | No. 10 | October 2013

Downloaded from https://academic.oup.com/annonc/article-abstract/24/10/2471/177745 by guest on 26 July 2018

editorials
Annals of Oncology
review of 321 randomized oncology trials, noted that these trials have become larger with time, and more likely to have corporate sponsorship, and that for-profit sponsorship was independently associated with endorsement of the experimental arm.

Many randomized trials also have the problem that the large numbers increase the statistical power of the study, and hence the study is more likely to be positive. While increased statistical power may be regarded as a virtue, such large trials are plagued by at least three major shortcomings. The first problem is that, if statistical significance is achieved, there is then often the assumption that reaching significance implies that the results are clinically important, but this may not be the case. For example, the randomized trial exploring the use of erlotinib in pancreatic cancer (that led to its approval in the United States) demonstrated that survival was improved statistically, with a hazard ratio of 0.82 for the gemcitabine plus erlotinib arm versus the gemcitabine arm (and with 1-year survival of 23% versus 17%), but the median survival only increased by 11 days [14]. Hence, large, randomized studies encourage one to set the efficacy bar low and to aim for gains that are measurable but unimportant.

The second major problem is that randomized trials may disregard the impact of patient subsets. If the study is positive, then the new therapy is adopted for all subsequent patients of this type, ignoring the fact that the benefit may have been restricted to a relatively small subpopulation, with potential harm in other subpopulations, and with neither benefit nor harm in others. Conversely, if the study is negative, the therapy may be discarded despite being of marked benefit in one or more distinct subpopulations.

The third problem is that large, randomized trials consume considerable patient and financial resources, so these are not available to assess other therapies or approaches. While it has been contended that ethically questionable randomized trials may be justified if they help payors decide whether or not to pay for a new therapy [4], we would argue that this would constitute a squandering of scarce research resources. In our opinion, it is much more important to use these resources to discover and develop new therapies for untreatable illnesses than it is to make decisions regarding payment if those decisions require trials beyond the point where the consensus of the expert medical community deems the treatment beneficial.

In conclusion, the rationale [4, 15] for abandoning equipoise as a criterion for determining whether a randomized, clinical trial is justified to evaluate new treatments and/or when such a trial should be terminated and, most significantly, when patients should be made cognizant of the disparities between arms, is faulty. A new standard for equipoise, based on the needs of regulatory agencies or payors for population-based data, is plagued by the financial incentive that payors have to delay approval for potentially expensive agents. Furthermore, the argument that patients with terminal diseases are not wronged by being denied a ‘partially evaluated’ treatment because they are not worse off than they would be outside of the trial is disingenuous at best [4]. They are wronged because they are denied a treatment that is widely believed to be lifesaving, and because there is sufficient evidence that, above a certain threshold value, the value of such treatments can be fairly certain [6, 16].

Concerns about enrollment should not mitigate the need to fully enlighten patients regarding the disproportion in treatment response between the two arms of a randomized trial. Undoubtedly such information, if presented in its entirety and transparently, would lead to inability to accrue well and/or early study termination, if one arm was widely believed to be better. It is our experience that very few patients would be willing to stay on the ineffective or marginally effective control arm of a trial just, so that payors can obtain information. Moreover, there is evidence that such strategies with dramatically effective drugs are costly, and that the lives lost due to delayed approvals makes this approach potentially harmful to the population at large.

As an alternative, it has been proposed that, for life-threatening diseases, regulatory agencies should set flexible standards that allow rapid approval of new agents [17]. Chabner [17] posits, and we agree, that high response rates (>50%), high disease-control rates (>75%), and an acceptable side effect profile in a biomarker-defined population of 75–100 subjects should be sufficient for accelerated approval.

In summary, there is no approach that completely eliminates the potential for error in drug approval. However, it appears that, with remarkably effective drugs for life-threatening illness, strategies that endorse randomization beyond equipoise are neither cost effective nor beneficial to individual patients or to society, and may at times be harmful to the population, and promote deception of trial participants. The equipoise standard for determining the ethics of randomized trials has served well for several decades. The advent of new drugs with remarkable efficacy and limited toxicity for patients with terminal cancer should not prompt abandoning this standard, so that randomized trials of these drugs can remain justified. Instead, for highly effective drugs, it would be useful to develop new study designs that prove efficacy rapidly, and consensus criteria that determine threshold values for salutary effects beyond which a randomized trial is not necessary.

R. Kurzrock1,* & D. J. Stewart2

1Division of Hematology and Oncology, University of California Moores Cancer Center, San Diego, USA
2Division of Medical Oncology, University of Ottawa, Ottawa, Canada

(*E-mail: rkurzrock@ucsd.edu)

disclosure

The authors have declared no conflicts of interest in regard to this manuscript.

references


