ADVANCES IN DRUG DEVELOPMENT

Current Developments in Oncology Drug Research

Section Editor: Mark J. Ratain, MD

Improving the Design of Pivotal Clinical Trials

Vinay K. Prasad, MD, MPH
Associate Professor of Medicine
Assistant Program Director
Hematology Oncology Fellowship Program
Oregon Health and Science University
Portland, Oregon

H&O What are some typical flaws in the design of pivotal clinical trials?

VP There are 5 common flaws in the design of pivotal clinical trials used to support the approval of cancer drugs by the US Food and Drug Administration (FDA). The first is an inappropriate or problematic endpoint. Many clinical trials study the effect of therapy on a surrogate endpoint, which is not a measure of how patients feel or function, but a stand-in variable. An example would be the measurement of tumor shrinkage or growth on a scan. The trial design then sets an arbitrary change for the surrogate. It is not just that the tumor shrank; it shrank beyond some arbitrary percentage. Carefully done studies have shown that many of these surrogate endpoints correlate poorly to subsequent changes in survival or quality of life.

Another concern is the treatment selected for the control arm. A randomized controlled trial of a treatment for cancer should test the new drug against either what doctors are already doing or the best available therapy. In the United States, the control arm typically consists of the best available therapy. However, clinical trials for US drug approvals are often run in global settings, where resource scarcity may mean that the control arm consists of an older inferior option. That is known as the straw man control, and renders trials useless to inform US practice.

Another potential flaw is for the effect size to be too small. For example, a randomized trial with a good control arm might have overall survival as an endpoint. However, the trial may be designed to detect a difference in survival that is statistically significant but not clinically meaningful. There may be slivers in Kaplan-Meier survival curves indicating trivial differences that are not meaningful to patients.

Postprotocol therapy is another concern. In trials run in global settings with scarce resources, patients with progressive disease may receive subsequent therapies that do not meet US standards. It may be possible to show that a new drug is better than an older drug in that setting, but the trial is not addressing whether the new drug is better than the older drug in the United States, where postprotocol therapy would be better.

Crossover is misused in 2 ways. For drugs that have unproven efficacy, crossover makes an overall survival endpoint uninterpretable, and this is problematic. In trials evaluating whether a drug can be moved forward in the treatment course, crossover must occur, to ask whether the drug administered earlier is superior to the standard of care of giving it later. Alyson Haslam, PhD, and I discuss crossover in an article appearing in the *Annals of Oncology*.

H&O What are the ramifications of these flaws?

VP Individually and collectively, these flaws make clinical trial data cloudy. They obscure information from patients and doctors. A trial may show that a new drug improves a surrogate endpoint, but not whether it helps patients live longer or better. A drug may improve survival for approximately 10 days, such as erlotinib (Tarceva, Genentech/Astellas) in pancreatic cancer, but this improvement is not clinically meaningful. Or a trial might show that a new drug is better than an inadequate comparator. An example would be trials showing that the combination

of ibrutinib (Imbruvica, Pharmacyclics/Janssen) and rituximab (Rituxan, Genentech/Biogen) is superior to rituximab in Waldenström macroglobulinemia. But is the ibrutinib/rituximab regimen superior to the combination chemoimmunotherapy regimens that doctors actually use? The answer is not known. The flaws in clinical trial design have a real impact, as they can generate poor information that is then used to guide treatment decisions.

> Clinical trials tend to enroll patients who are almost a decade younger than the average cancer patient in the United States.

H&O Are there any principles that can be applied to the design of clinical trials to improve their quality?

VP In oncology, the optimal design is most likely a randomized clinical trial that tests the new therapy against the best treatment available in the United States. Most of these trials are evaluating a drug to obtain approval from the FDA, and the most straightforward way to do this is to run the trial in the United States and test it against the best available therapy.

H&O How can a doctor assess the quality of a clinical trial when considering enrollment of a patient?

VP The most important information to consider at the outset is the extent of activity shown by the experimental drug as a single agent in the disease. Points in a trial's favor include if it is randomized, run by a cooperative group, and non-conflicted, meaning with no influence from the manufacturer of the drug tested, which often stands to make a windfall, and no influence from physicians who are receiving consulting payments from the manufacturer. When a doctor is considering enrolling a patient in a nonrandomized trial, he or she should consider whether the patient has exhausted all of the proven options. It is difficult to know when to enroll a patient, and doctors should individualize the decision.

H&O How can the interpretation or presentation of trial results misrepresent the data?

VP There are 3 major ways. A report can spin borderline or negative findings so that they appear positive. This type of spin is prevalent in the reporting of many clinical trials in oncology. There is selective reporting, in that negative trials are less likely to be published, and more likely to be published after a delay. Another way involves what I call Twitter cheerleaders. People are cheerleading for new products based on limited information. A classic example is that a drug company issues a press release saying that a drug is beneficial—with a significant *P* value—but without specifying the magnitude of benefit. Then someone on Twitter says that the drug is practice-changing.

H&O Do results from clinical trials differ from real-world experience?

VP They do. Clinical trials tend to enroll patients who are almost a decade younger than the average cancer patient in the United States. Enrolled patients also have fewer comorbidities. In a younger patient population with fewer comorbidities, a very toxic drug can be pushed to a high dose and achieve a marginal benefit. When the drug is used in the real world, in older and frailer patients, those small benefits may evaporate or there may even be net harm. One example is the use of sorafenib (Nexavar, Bayer) in metastatic liver cancer. In the pivotal trial that led to approval, the median survival was 11 months with the drug vs 8 months in the control arm. In a real-world analysis of Medicare patients by Sanoff and colleagues, those treated with sorafenib lived approximately 2 to 3 months. Those who were not treated with sorafenib but were matched to treated patients via a propensity score also lived approximately 2 to 3 months. Therefore, this drug with a marginal benefit in an ideal population had less benefit, or none at all, in the real world. The study showed that patients treated with sorafenib lived about half as long as patients treated with placebo in the clinical study. This finding highlights the selection bias in the trial.

H&O Can the FDA improve the drug approval process?

VP When considering which drugs to approve, the FDA should use surrogate endpoints sparingly. My colleagues and I have performed several studies showing that two-thirds of drug approvals were based on surrogate endpoints.

The FDA has the legal and statutory authority to set minimum benefits. The FDA does not have to approve a drug based on a small percent improvement in a surrogate endpoint. This idea was proposed by Bekelman and Joffe in *JAMA*.

H&O What are the appropriate circumstances for an accelerated approval?

VP With accelerated approval, a drug is provisionally approved based on a surrogate endpoint and later has a postmarketing commitment. To utilize this pathway, the new indication must be an unmet need. Research by my colleagues and I have shown that the phrase "unmet medical need" is used indiscriminately. It is used in settings that are dire, as well as in those with a favorable 5-year survival. The phrase can encompass conditions that are rare or common, and those with few or many treatment options. I believe that the term "unmet medical need" should be reserved for those conditions that are dire and rare, and that have few treatment options. Accelerated approval should be used sparingly, and it is necessary to enforce the initiation of postmarketing studies.

H&O What is the role of postmarketing efficacy studies?

VP Ideally, postmarketing efficacy studies could be transformational and answer unresolved questions. Unfortunately, some lapses exist in regulatory policy concerning these studies. In an analysis of 34 postmarketing studies by Zettler and Nabhan, 44% had been completed, 41% were ongoing, 9% were pending, and 6% were terminated.

A publication by the FDA found that many accelerated approvals are later converted to full approvals. In many cases, however, both the accelerated approval and the full approval were based on a surrogate endpoint (either the same or a different one). If the initial approval is based on a surrogate endpoint, then the full approval should measure survival or quality of life. Many drugs therefore come to the market without any proof that they improve survival or quality of life, and they remain on the market for many years without that proof. This is problematic.

H&O Do you foresee any innovations that could improve clinical trial design?

VP There is increasing recognition of the problems with the current design of clinical trials. Researchers in this field believe that the system is due for reform, in both the United States and Europe. Many of the problems seen with US studies have also been identified in Europe, suggesting that there are global problems in trial design and drug development. The fundamental problem is that more drugs can be approved if the regulatory bar for approval is lowered. There will be more drug approvals, but also more uncertainty regarding whether the treatments actually improve outcome. Patients do not just want more options, they want good options. For the options to be good ones, it will be necessary to enforce some standards of efficacy to show that the drugs help patients live longer and better. It is especially important to do this with therapies, like cancer drugs, that are very toxic and expensive. The status quo is untenable.

Disclosure

Dr Prasad reports receiving royalties from his book Ending Medical Reversal*; that his work is funded by the Laura and John Arnold Foundation; that he has received honoraria for Grand Rounds/lectures from several universities, medical centers, and professional societies; and that he has received payments for contributions to Medscape.*

Suggested Readings

Beaver JA, Howie LJ, Pelosof L, et al. A 25-year experience of US Food and Drug Administration accelerated approval of malignant hematology and oncology drugs and biologics: a review. *JAMA Oncol*. 2018;4(6):849-856.

Bekelman JE, Joffe S. Three steps toward a more sustainable path for targeted cancer drugs. *JAMA*. 2018;319(21):2167-2168.

Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13 [published online October 4, 2017]. *BMJ*. 2017;359:j4530. doi:10.1136/bmj.j4530.

Haslam A, Prasad V. When is crossover desirable in cancer drug trials and when is it problematic? *Ann Oncol*. 2018;29(5):1079-1081.

Kemp R, Prasad V. Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused? *BMC Med*. 2017;15(1):134.

Kim C, Prasad V. Cancer drugs approved on the basis of a surrogate end point and subsequent overall survival: an analysis of 5 years of US Food and Drug Administration approvals. *JAMA Intern Med*. 2015;175(12):1992-1994.

Kim C, Prasad V. Strength of validation for surrogate end points used in the US Food and Drug Administration's approval of oncology drugs. *Mayo Clin Proc*. 2016;91(6):713-725.

Lu E, Shatzel J, Shin F, Prasad V. What constitutes an "unmet medical need" in oncology? An empirical evaluation of author usage in the biomedical literature. *Semin Oncol*. 2017;44(1):8-12.

Prasad V, Kim C, Burotto M, Vandross A. The strength of association between surrogate end points and survival in oncology: a systematic review of trial-level meta-analyses. *JAMA Intern Med*. 2015;175(8):1389-1398.

Sanoff HK, Chang Y, Lund JL, O'Neil BH, Dusetzina SB. Immortal time bias or sorafenib effect in elderly patients with HCC? *Hepatology*. 2017;66(2):678-679.

Sanoff HK, Chang Y, Lund JL, O'Neil BH, Dusetzina SB. Sorafenib effectiveness in advanced hepatocellular carcinoma. *Oncologist*. 2016;21(9):1113-1120.

Zettler M, Nabhan C. Fulfillment of postmarketing requirements to the FDA for therapies granted oncology indications between 2011 and 2016. *JAMA Oncol*. 2018;4(7):993-994.