An Essay on Rearranging the Deck Chairs: What’s Wrong with the Cancer Trials System?

Commentary on Gray et al., p. 1966

Derek Raghavan

In 2006, >1 million new patients with cancer will be diagnosed in the United States and <5% of these will be entered into structured clinical trials. Regrettably, many of these studies will be focused on approval processes for pharmaceutical industry marketing rather than on innovation and the provision of solutions to the problem of intractable malignancy. Although the concepts advanced in the article of the investigators of the Eastern Cooperative Oncology Group (1) are laudable and important, their approach does not really target the critical issues (i.e., lack of accrual to trials and the problems of the regulatory processes governing trial design and implementation among NIH-supported research groups).

To focus initially on the work of Gray et al. (1), it is evident that these experienced investigators have identified clearly the problems of accrual, compounded by the panoply of novel compounds that now require testing in clinical trials. They have attempted to identify workable strategies to deal with this dichotomy. Certainly their proposals are reasonable, for example, the use of randomized phase II design, randomized discontinuation studies, factorial designs, and approaches predicated on Bayesian analysis.

The authors suggest an approach of fast-tracking using randomization to reduce sample differences between study populations and testing combinations of novel compounds in an attempt to move quickly toward a quantum leap forward. The problem is that each approach compounds the felony imposed by small case sample size, imprecise or novel end points for which long-term biological significance is unclear, and in which treatment interactions have not been defined.

The concept that randomized phase II design will overcome the differences in the characteristics of sequentially sampled populations is theoretically attractive. There really is no evidence, however, that this is true when the sample size is as small as required for a standard phase II trial, especially one with early stopping rules. When two or more phase II studies are compared synchronously, hoping that randomization will overcome population differences, there is implicit acceptance of significant δ and ε errors (2). This seems problematic when it is time to decide which agent moves to the next step of testing. For example, in the setting in which patients with hormone-refractory prostate cancer from academic centers are sent for randomized phase II testing, groups of 40 to 50 patients are not likely to overcome the variability of demographic traits, multiple prior sequences of hormonal and biological medications, prior bone-stabilizing drugs, and other supportive treatments. It should not be forgotten that such centers use patient resources as assiduously as possible; thus, any number of permutations of prior treatment interactions and late consequences is possible, which could confound ultimate assessment.

Similarly, the randomized discontinuation design is problematic. We already know that the assessment of response is highly imperfect and that our approach to the assessment of “stability” or “failure to progress” is even less perfect. The studies from Warr et al. (3) two decades ago showed clearly that our clinical and radiologic approaches to measurement of small changes are subject to wide subjective error; this really has not changed. This is likely to be compounded in the ongoing assessment of this end point for many of the new, targeted therapies. Unless there is a dramatic difference in time to progression, as has been seen in some of the newer therapies for renal carcinoma, this approach will fail because there are many potential confounding variables in assessment of time to progression as stated by Gray et al. (1). These really are not overcome by randomized discontinuation. Why is this therefore likely to be a productive approach unless the agent represents a true breakthrough (which would be obvious in much simpler clinical trial designs)? The issue becomes most vexing when this model is considered for multiple arm comparisons.

Factorial designs are a particularly appealing way, in theory, to maximize the use of a limited number of cases. Thus, an initial randomization compares two treatment approaches, and then the patients are subjected to a second randomization, thus averaging out the differences between the cases being compared in each echelon. For this to be reproducibly accurate, however, the concept requires an absence of interaction between the various treatments. Given the lack of clear understanding of the multiple interactions and downstream effects of the new, targeted agents, the use of this approach to test combination regimens is likely to be fraught with unpredictable hazard, with occult interactions that may never be identified in economical clinical trials with restricted case sample size.

Finally, a detailed discussion of the problems of Bayesian analyses to determine early stopping rules is beyond the scope of this brief editorial. There is, however, an extensive literature on the limitations of application of historical controls and repeated measurements of samples and outcomes in small populations, and these issues are not simply overcome by the elegance of Bayesian logic.

What is the solution then? Perhaps it will require government, which casts such a long shadow in the health care industry, to be more responsible and to consider carefully the nature of the problem to which it contributes. The use of
government imperatives to minimize the current policy of reimbursement for established ineffective treatments for some relapsed, metastatic cancers could reduce the wastage of clinical resources and the profligate use of the time of dying patients, could give patients greater access and incentive for participation in well-designed studies, and might allow improved patient outcome via the structured use of novel, promising therapies. The provision of broader levels of fiscal or infrastructural support (offset by a reduction of wasted expenditures) to allow patients to overcome some of the rigors of entry into clinical trials (fiscal burdens, requirements for transport and other support, and access to clinician time) would also help. At a much simpler level, the government and its health regulators might even try to assess and understand why so few formal surgical trials exist when surgeons remain the gatekeepers of clinical cancer practice.

Another focus for government could be the process of protocol and program review. In the United States, there is an ever-present potential for conflict of interest because cooperative group protocols are reviewed by members of competing groups and/or by government personnel with a range of potential occult conflicts and external pressures. Government functionaries, mindful of the level of available resources, make pragmatic decisions regarding cooperative group oversight that often seem to be based more often on resource availability than purely on science or the potential for clinical benefit. Members of study sections, aware of the diminishing level of clinical and fiscal resources, similarly may make judgments potentially marred by subconscious conflicts of interest.

One creative solution could be to learn from our colleagues in Europe. The European Organization for Research and Treatment of Cancer and Cancer Research UK both use panels of extramural, international experts to assist in the evaluation of their programs and protocols. This avoids the problem of a decision process potentially confounded by self-interest.

A serious attempt to address some of these issues was the creation of the Clinical Trials Working Group of the National Cancer Advisory Board, an appointed panel of 38 experts given the charge of revising the approach to clinical trials of the National Cancer Institute (4). After a lengthy period of deliberation, an extensive menu for change was created that included (a) improved coordination and collaboration among the elements of the current system, including industry and regulatory agencies; (b) improved design and prioritization of trials; (c) improved process, including design, data capture, data sharing, and infrastructure; and (d) improved operational efficiency with increased patient accrual and reduced operational barriers.

Although important and meritorious, this hallmark document was also problematic. It is puzzling that an agency seeking review of its processes and the provision of novel solutions selected a panel with >50% of its membership drawn from its own ranks under the chairmanship of its own leaders. Notwithstanding the excellence and reputation of many key government medical officials, they are nonetheless subject to the whim of politicians and have reporting relationships within government that are hardly unfettered. This seems to have resulted, yet again, in the traditional need for governmental control and direction of the range of proposed trial functions, which seemed to dominate the agenda of proposed solutions. Once again, no attempt was made to embed a review process with input drawn from outside the system. For example, the various proposed new regulatory and consultative committees were all to be predominantly composed of government medical officers or investigators funded by National Cancer Institute system (and thus beholden to it), with only a smattering of truly external members.

The greatest problem, however, was that the manifesto ignored the fundamental problem. The majority of physicians who treat cancer and the majority of cancer patients are completely uninvolved in the government trial process. The recommendations were aimed more at the small proportion of those who are currently within the system. It was simply naïve to consider that industry would be influenced globally by a series of peripheral measures that focused on the common pathways of drug assessment within the ambit of the National Cancer Institute and the Food and Drug Administration. Similarly, it was unrealistic to believe that mechanisms of data capture and data sharing, accreditation of investigators, and simplification of access to complex trials would actually target the dominant population of combatants who are uninvolved in the government’s war on cancer, which is scheduled to end in 2015. The isolation of the beltway seems to have dominated the scope of the debate and thus avoided the key issues of reimbursement for relatively ineffective therapies and of patient and investigator recruitment in the real world.

Perhaps it is time for government, a major stakeholder in any health system, to think outside the box. Instead of rearranging the deck chairs on a sinking ship, it may be time to chart a course to avoid the icebergs.

References
An Essay on Rearranging the Deck Chairs: What's Wrong with the Cancer Trials System?

Derek Raghavan


Updated version  Access the most recent version of this article at:  
[http://clincancerres.aacrjournals.org/content/12/7/1949](http://clincancerres.aacrjournals.org/content/12/7/1949)

Cited articles  This article cites 3 articles, 2 of which you can access for free at:  
[http://clincancerres.aacrjournals.org/content/12/7/1949.full.html#ref-list-1](http://clincancerres.aacrjournals.org/content/12/7/1949.full.html#ref-list-1)

Citing articles  This article has been cited by 3 HighWire-hosted articles. Access the articles at:  
[http://clincancerres.aacrjournals.org/content/12/7/1949.full.html#related-urls](http://clincancerres.aacrjournals.org/content/12/7/1949.full.html#related-urls)

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.