I-SPY 2: A Glimpse into the Future of Phase II drug Development?

David Harrington, Ph.D.

Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute

Department of Biostatistics, Harvard T.H. Chan School of Public Health

Giovanni Parmigiani , Ph.D.

Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute

Department of Biostatistics, Harvard T.H. Chan School of Public Health

The papers by Rugo and Park in this issue report results from the I-SPY-2 platform, a promising adaptive strategy for matching targeted therapies for breast cancer with patients most likely to benefit. I-SPY-2 identified two therapies (velaparib with carboplatin in triple negative breast cancer and neratinib in HER2+/HR- cancer) that met pre-specified criteria for testing in larger, phase III trials. The value of ISPY-2, however, may well go beyond the clinical results described here. Adaptive multi-arm trials such as I-SPY 2 have the potential to answer several questions simultaneously and more efficiently than traditional trial designs. Which of several promising therapies appear best suited for larger, confirmatory trials? Which patients should be asked to participate in those trials? Is the chance of success in subsequent larger trials sufficient to justify the expense and time needed?

Therapies designed to target molecular subtypes of cancer may increase the chances of good responses and, equally importantly, may be useful in avoiding treatments in settings where meaningful benefit is unlikely. The challenges, however, in identifying successful targeted therapies in cancer are substantial. Targeted therapies may fail to hit their target, they may not have the predicted effect when they do, and they also may have a positive effect in the absence of a recognized target. Traditional phase II designs that test treatments one at the time in heterogeneous groups of patients have created a traffic jam – there are too many new drugs, and the signal for a treatment effect can be diluted in a heterogeneous mix of trial participants[[1]](#endnote-1). The US FDA[[2]](#endnote-2) and the European Medicines Agency[[3]](#endnote-3) have acknowledged that the commonly used designs need a makeover.

The efficiency of multi-arm early phase trials has long been recognized, but I-SPY 2 differs in important ways from traditional early phase trials. Since March 2010, the I-SPY 2 platform has been used to compare 5 experimental therapies with a common control in breast cancer subsets with ten distinct biomarker signatures. The randomization is stratified (8 strata defined by HER2, hormone receptor status and MammaPrint classification), and adaptive randomization is used within strata to increase the likelihood of assignment to a therapy that accrues evidence of being more successful than its control at inducing pathologic complete responses in patients with locally advanced cancers. New drugs can enter the platform as they emerge from phase I testing, and exit the platform with an estimate of the chances of future success in a phase III trial of pre-specified size. The platform may be an appealing setting for cooperation among pharmaceutical companies and academic investigators. The entire process, including design and analyses, are carried out dynamically using Bayesian methodologies.

Oncology has been slow to adopt Bayesian designs even though they are often well suited to settings where inference and decisions benefit from adaptation on accruing information. Some of the reluctance stems from a natural discomfort with replacing a familiar approach that has provided some success in the past. There are other, more substantive reasons to be cautious about this new path. ISPY-2 was designed in 2009. In the world of trial design, it is still in its infancy. There is much to be learned about the statistical models used to adaptively adjust randomization fractions and to predict the chances of success in a future trial. How robust are the adaptive randomization probabilities and the predictive probabilities of success in a phase 3 trial to misspecifications of the model? What are effective ways to communicate to our clinical colleagues the modeling assumptions used, the potential vulnerability of the model to errors, and the best ways to explain these designs to trial participants? What visual and numerical summaries provide insight into the trial data? Simple summary statistics such as odds ratios or relative risk can be misleading, and the usual CONSORT diagram does not reflect the dynamics of the I-SPY 2 randomization. How will the predicted high chance of future success (80%) upset equipoise for trial investigators or influence the kinds of patients they choose to enroll or not enroll in a future trial?

It is important to investigate these questions in depth. An important first step would be the availability of open source software for managing the randomization and analysis steps of a platform like I-SPY 2. Nonetheless, I-SPY 2 is an important addition to the inventory of trial designs. The two figures 2 in both papers showing estimated distributions of pathologic complete response rates are appealing. Ninety-five percent probability intervals are provided in the paper, but the graphs in figure 2 make it easy to identify, for instance, 90% or 99% intervals. Clinicians can interpret the results of the trial consistent with their own sense of acceptable uncertainty. Adaptively adjusting randomization probabilities makes much more sense than specifying an unbalanced but fixed randomization at the beginning of a trial. Perhaps most importantly, I-SPY 2 holistically integrates the ideas of Bayesian design and analysis in the important setting of phase 2 testing of new cancer drugs. The design of the platform acknowledges the complexity of phase II testing in cancer.

As George Box famously wrote, “Essentially, all [statistical] models are wrong, but some are useful.[[4]](#endnote-4)” The most useful models add much more than just statistical information; they help bring clarity to settings where noise and uncertainty threaten to overwhelm progress. The fundamental tenets behind the I-SPY 2 platform and model – the multi-arm platform, options for graduation and addition of drugs, adaptive randomization, and prediction of success in confirmatory trials – are important first steps toward efficient use of clinical resources. As more new targets and drugs are discovered, traditional statistical designs, at best cumbersome and inefficient today, will be wholly insufficient for matching patients with effective drugs. We applaud the use of I-SPY 2 described here, and urge continued innovation in trial design, especially in both earlier phase 1 and later phase 3 settings.

1. Betensky, RA, Louis, DN and Cairncross, JG. *Influence of unrecognized molecular heterogeneity on randomized clinical trials.* J Clin Oncol. 2002 May 15;20(10):2495-9. [↑](#endnote-ref-1)
2. FDA. *Guidance for Industry. Adaptive Design Clinical Trials for Drugs and Biologics* [online], http://www.fda.gov/downloads/DrugsGuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf (2010). [↑](#endnote-ref-2)
3. Committee For Medicinal Products For Human Use (CHMP). *Reflection Paper On Methodological Issues In Confirmatory Clinical Trials Planned With An Adaptive Design* [online], http://www.ema.europa.eu/docs/en\_GB/document\_library/Scientific\_guideline/2009/09/WC500003616.pdf (2007). [↑](#endnote-ref-3)
4. Box, G. E. P., and Draper, N. R., (1987), *Empirical Model Building and Response Surfaces*, John Wiley & Sons, New York, NY, p 424. [↑](#endnote-ref-4)