Statistical Challenges with the Advances in Cancer Therapies

Rajeshwari Sridhara

Abstract Statistical challenges in designing, analyzing and interpreting the data are being encountered with the recent development of new classes of drugs to treat cancer. The existing paradigm of drug development from Phase I to Phase III clinical trials is not optimal. New and innovative trial designs and statistical methods are needed to evaluate the new classes of drugs. In this chapter we present the regulatory considerations in the evaluation of drug products, the drug development paradigm in the last century and the current time, and the statistical challenges that need to be addressed.

keywords Regulations · Cancer drug development paradigm · Immunotherapies

1 Regulatory Considerations

With the signing into law of the Kefauver-Harris Drug Amendments to the Food and Drug Cosmetic Act in 1962, drug manufacturers were for the first time required to prove to the US FDA the effectiveness of their products before marketing them [1]. This amendment was intended to ensure both drug efficacy and safety, and gave a statistical framework for conducting clinical trials to prove the effectiveness of drug products. Section 505(d) of the Food and Drug Cosmetic Act [2, 3] as amended states that "…evidence consisting of adequate and well-controlled investigations, including clinical investigations, by qualified scientific experts, that proves the drug will have the effect claimed by its labeling …". This statement has been used as the regulatory standard for establishing evidence and interpreted to mean the following: the evidence should be reproduced in at least two independent studies, the probability of one-sided type I error should be controlled at a threshold of 0.025, a clinically meaningful treatment effect should in general be established

R. Sridhara (✉)
Office of Biostatistics, Center of Drug Evaluation and Research,
US Food and Drug Administration, Silver Spring, MD 20993, USA
e-mail: Rajeshwari.Sridhara@fda.hhs.gov

S. Matsui and J. Crowley (eds.), *Frontiers of Biostatistical Methods and Applications in Clinical Oncology*, DOI 10.1007/978-981-10-0126-0_2

even if the results are statistically significant, and the product should have an acceptable risk-benefit profile.

Two decades later, in 1981, the FDA and the Department of Health and Human Services revised the regulations for the protection of human subjects, detailing the contents of informed consent and widening the representation in institutional review boards. Another landmark in the history of the FDA was the publication of regulations in 1991 establishing a new path to accelerate the review of drugs for life-threatening diseases. Today we have two regulatory pathways for marketing approval of drug products: regular or traditional approval and accelerated approval.

The regular approval decision is based on demonstrated clinical benefit of the drug product, for example, improved overall survival in cancer patients compared to placebo, or on an outcome that clearly benefits a patient, such as an improvement in disease related symptoms. The accelerated approval decision is based on a surrogate endpoint reasonably likely to predict clinical benefit, such as objective tumor response rate, and the treatment effect should be better than available therapy. Products approved under the accelerated approval pathway are, however, required to subsequently establish improved clinical benefit by conducting a confirmatory clinical trial.

The statistical considerations in evaluating drug products include (1) quality and quantity of data, (2) design of the study, (3) method of analyses, and (4) interpretation of the results from the analyses. With respect to clinical trial design, the important considerations are whether the study is randomized or not, the presence or absence of adaptive features, whether a superiority or non-inferiority hypothesis is tested, the extent to which the overall false positive rate is controlled, and whether the results are replicated. Important considerations in the analyses include clear definition, measurement and validation of the outcome of interest; the statistic used to test the hypothesis and whether the data conform to the assumptions of the chosen analysis method; whether any subgroups were identified and pre-specified to be tested; imbalances between treatment groups in the subgroup; and finally whether multiple hypothesis testing was conducted.

2 The Drug Development Paradigm in the Last Century

The development of cytotoxic drugs, the predominant treatment of cancer in the last century, has generally been comprised of a step-wise approach with clearly defined phases of clinical trials: Phase I trials for dose finding, Phase II trials to determine drug activity, and Phase III trials for confirming efficacy. Phase I trials have been designed to find the maximum tolerated dose (MTD), commonly using an algorithmic design such as a 3 + 3 design or more recently a model-based design (for example, modified continual reassessment methodology). In these trials, the dose was continuously increased until dose limiting toxicity (DLT) was observed. A lower dose than the DLT dose was considered the maximum tolerated dose (MTD). The MTD was further evaluated in the next phases of the study to assess

the efficacy and overall risk-benefit of the drug. In this cytotoxic paradigm a 'more is better' approach was used, because of the desire to kill the maximum number of cancer cells. For cytotoxic therapies, there were reasonably good preclinical models prior to conducting first-in-human Phase I studies, treatment was limited to a finite number of treatment cycles of 21–28 days, the dose given to a patient was based on body surface area, toxicities were observed in a short period of time, and the toxicities (hematologic, neurologic, etc.) were well characterized.

The Phase II single-arm trials evaluated activity of the drug using intermediate outcomes such as tumor response rate that could be observed in a relatively short time. Typically, these trials were designed using the Simon two-stage approach [4] as single-arm studies. In this approach, patients would be enrolled and treated in two stages. If the tumor response rate in the group of patients enrolled and treated at MTD in the first stage was less than a pre-specified threshold, the drug would not be studied any further; and if response rate was more than this threshold, an additional group of patients would be enrolled to the second stage. Only if the overall response rate was more than a desired threshold in the two groups of patients combined would the drug would be further evaluated in Phase III trials.

The confirmatory Phase III trials evaluating the efficacy and safety of the drug were randomized controlled trials comparing the investigational drug to the standard of care, with overall survival as the primary outcome of the clinical trial. Because the toxicities were well characterized for the cytotoxic products and the treatment was limited to a finite number of treatment cycles, the toxicities observed during the different phases of drug development formed an adequate basis to guide physicians in the management of patient treatment.

3 The Current Drug Development Paradigm

With the understanding of the biology of the disease and the development of non-cytotoxic drugs such has kinase inhibitors and immunotherapy, cancer treatment options have changed in the last two decades. In terms of both toxicity and activity/efficacy, these products are very different from cytotoxic products. There are few if any good pre-clinical models to predict the likely starting dose and toxicities in humans, although these products are in general better tolerated. Severe toxicities of these drugs are not always observed in a short duration of time, and treatment is not limited to a few cycles, but typically continued until disease progression is observed. Many of these products are taken orally and administered in fixed doses rather than based on body surface area. Often a long-term effect on overall survival is observed in the absence of objective tumor response rate (example: sorafenib, ipilimumab) [5, 6]. Thus, the cytotoxic paradigm fails in every phase of drug development for the current generation of drug products. The cytotoxicity-based definition of dose-limiting toxicity is no longer useful, because many of these products do not have the well characterized hematologic or non-hematologic toxicities. For example, the kinase inhibitor erlotinib has severe

skin toxicity, which is not observed with typical cytotoxic drugs. Many of the toxicities do not occur within the short time of observation in the Phase I trials where more refractory patients with a shorter life expectancy are enrolled. Some of these drugs may not shrink tumors but rather stabilize the disease, resulting in poor response rates and requiring randomized Phase II studies to better understand the activity of the products with respect to other outcomes such as progression-free survival. Because of the unknown long-term toxicities of these drugs it is not uncommon to have dose interruptions and reductions in Phase III trials, with the result that when the confirmatory clinical trial is completed, recommended dose and monitoring guidelines for patient care are not always clear.

3.1 Biomarker-Based Clinical Trials

The patient population enrolled in a clinical trial is recognized to be heterogeneous, (for example with respect to age, race, gender, genetic markers, subgroups of the disease, etc.), despite stringent inclusion and exclusion criteria. Therefore, when confirmatory clinical trial results do not demonstrate efficacy of the investigational drug, it is common to hypothesize that the drug is likely to be effective in a subgroup of the population. However, the challenge is in finding the specific subgroup that may benefit from the investigational drug. It is important to recognize whether the subgroup is defined based on a prognostic or predictive biomarker or both.

A prognostic biomarker is a biomarker that is measured at baseline (prior to administration of a treatment) that correlates with the treatment outcome for a heterogeneous set of patients and is independent of the treatment (Fig. 1a). For example, stage of disease that is measured at baseline is a prognostic marker of the overall survival of a given patient irrespective of the treatment received. A predictive marker is a biomarker that is measured at baseline prior to administration of a treatment that predicts whether a particular treatment is likely to be beneficial and it is associated with outcome of a specific therapy (Fig. 1b). Based on the predictive marker status, it is expected that there would be a differential benefit of a given treatment. For example, patients with metastatic melanoma with BRAF mutations benefit from BRAF inhibitors such as vemurafenib [7] and dabrafenib [8], and on the contrary, patients whose tumor is BRAF-negative (i.e., the BRAF gene is not mutated, or wild type) do not benefit from these treatments. Thus in many cases the biomarker status may guide the treatment options.

Various adaptive designs have been used and reported in the literature to identify and evaluate prognostic and predictive biomarkers. An ideal design would be to use a biomarker-stratified, randomized design as shown in Fig. 2. An example of this design is the lung cancer MARVEL trial [9] in which the patients' tumors were assessed prior to randomization for epidermal growth factor receptor gene (EGFR) status as measured by fluorescent in situ hybridization (FISH). Randomization was stratified by the EGFR status, and patients are randomly assigned to receive either

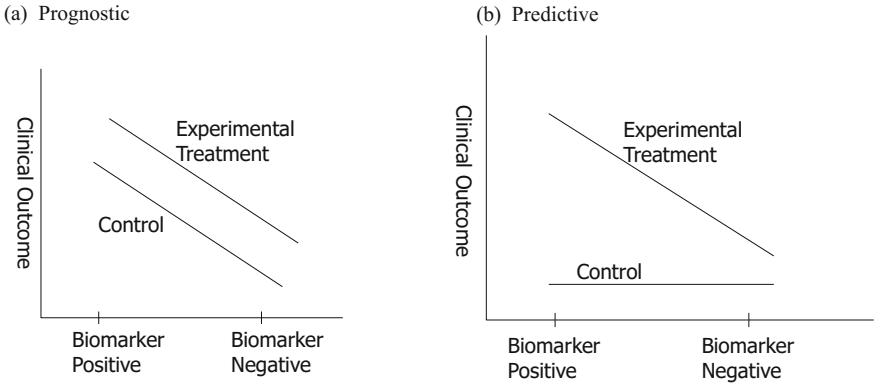

Fig. 1 Prognostic and predictive markers

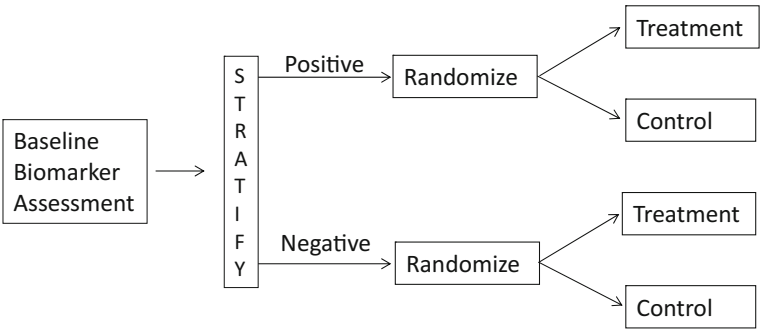

Fig. 2 Biomarker-stratified, randomized design

erlotinib or pemetrexed. In this design, the biomarker status is known for all randomized patients, and it can be evaluated as a prognostic and a predictive marker. On the other hand, if there is scientific evidence that given the mechanism of action of a particular drug it is unlikely that patients with biomarker-negative tumors would benefit from that drug, then an enrichment design (Fig. 3) is preferred as in the example of vemurafenib clinical trial where only patients whose tumor expressed BRAF mutation [7]. However, such a design assumes that the biomarker is predictive, and as such this design does not lend to evaluation of the biomarker as a prognostic or a predictive biomarker since marker-negative patients are not studied. Use of enrichment designs have increased with the development of targeted therapies. However designing such trials can be challenging as often the treatment effect of the standard of care in the enriched population may be unknown due to lack of information on the biomarker of interest in the historical control resulting in potentially underpowered Phase III studies, or the prevalence of the biomarker subgroup may be too small for a randomized clinical trial to be feasible.

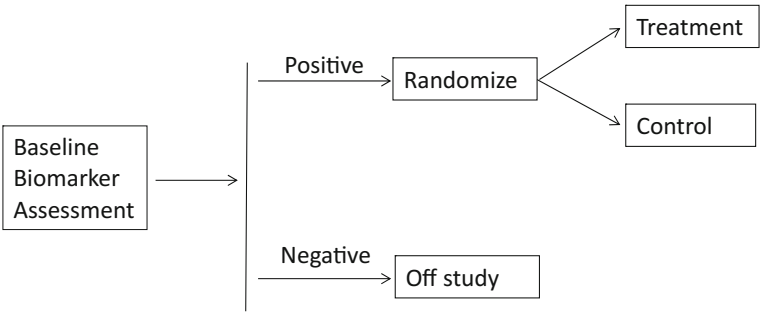

Fig. 3 Enrichment design

The ideal goal is to treat patients who benefit from a drug while not exposing patients who may not benefit and experience unwanted toxicity. However, due to the complex biology of the diseases not all characteristics that influence the outcome are measurable or known, and it is difficult to identify characteristics of patients who are likely to respond to a given treatment. Clinical trial designs have been proposed that evaluate predictive and prognostic molecular biomarkers and identify subgroups of patients who are likely to benefit from a given treatment after the clinical trial is completed in all patients [10–12].

More complex designs with adaptive enrichment strategies where enrichment occurs during the course of the clinical trial based on interim analysis of the data have also been suggested [13, 14]. Such designs with pre-planned decision criteria provide a scientific strategy to select the enriched population based on data accumulated in the initial stages of the clinical trial. Recently clinical trial designs [15–18] that can evaluate multiple diseases, multiple molecular biomarkers and/or multiple drugs (umbrella, platform, or basket trials) have been adopted in disease areas with unmet medical need. These trials typically have one umbrella or master protocol with a central governance structure, with adaptive features that allow adding and removing treatment arms, and are an efficient way of using patient resources. These clinical trials require adequate resources, coordination among different stakeholders and a trial network to conduct the studies. The approval of a new drug based on another trial while the current trial is ongoing, frequent adaptations to the design, multiple hypotheses testing, and overlapping characteristics of patients among two or more subgroups can pose challenges in execution and interpretation of the results of such clinical trials. Careful and detailed pre-planning, particularly in international studies, is essential.

Another component of the biomarker-based clinical trials is the companion or complementary diagnostics that are essential in defining the subgroups. Analytical validation of the biomarker assay based on performance (precision, accuracy, sensitivity and specificity), and quantitative and qualitative variability (e.g., differences in platforms, labs, technicians) are crucial in ensuring replication and interpretation of results. Because the use of a targeted drug is often tied to a diagnostic device in identifying the patient to be treated, there needs to be

coordination between the drug and device manufacturing companies as well as co-development of drug and device during the course of the product development cycle [19]. Often clinical trials are conducted with an investigator- or site-based diagnostic. It can be a challenge in evaluating the drug-device product for regulatory approval if the investigator- or site-based diagnostic differs with respect to operating characteristics from the scaled up version that is manufactured at a device manufacturing company.

3.2 Clinical Trials Evaluating Immunotherapy

Unlike chemotherapy and other targeted therapies, immunotherapy activates the immune system and thus indirectly targets the malignant disease. Thus, the early assessment of activity of products using tumor-based endpoints such as objective tumor response rate may not be ideal. Table 1 lists the FDA-approved immunotherapy products for the treatment of patients with advanced metastatic disease. These products have been approved under both accelerated approval and

Table 1 Immunotherapy products USFDA approved in metastatic diseases

Ipilimumab	Pembrolizumab	Nivolumab
March 2011, RA Unresectable/metastatic melanoma	**September 2014, AA** Unresectable/metastatic melanoma after Ipilimumab and BRAF inhibitor when indicated	**December 2014, AA** Unresectable/metastatic melanoma after Ipilimumab and BRAF inhibitor where indicated
	October 2015, AA PD-L1 + metastatic NSCLC after platinum based chemo	**March 2015, RA** Metastatic squamous NSCLC after platinum based chemotherapy
	December 2015, RA Unresectable or metastatic melanoma	**September 2015, RA** as single agent in unresectable/metastatic melanoma with BRAF wild type tumor **January 2016, AA** combination with Ipilimumab in unresectable/metastatic melanoma **AA** in BRAF mutant unresectable/ metastatic melanoma
		October 2015, RA metastatic NSCLC after platinum based chemo
		November 2015, RA metastatic RCC after anti-angiogenic treatment

regular approval provisions. The observed objective response rates were not always large, although duration of response tended to be long among those who have a response, and there were no meaningful differences observed in progression-free survival despite significant differences in overall survival [20]. In the immunotherapy clinical trials, it is also common to observe non-proportionality of hazard function in the analysis of progression-free survival [21]. Although some of the clinical trials evaluating antibodies blocking programmed cell death receptor 1 (PD-1) appear to suggest that programmed death ligand 1 (PD-L1) expression may be a predictive marker, it has not been evaluated systematically and it is unclear what threshold or cut-off value for PD-L1 expression is optimal in identifying the subgroup that benefits from these products [22]. In general in the clinical trials for these products the treatment continued until disease progression was observed. Because of this design, it is not known if the treatment can be stopped after a finite number of cycles of therapy or whether continued use is necessary. Although the currently approved products have demonstrated a favorable benefit-to-risk ratio, these early approvals have relatively short follow-up, and the safety of long-term use of these products is unknown at this time.

In designing future studies, the challenges will be in selecting the optimal endpoints for evaluation of these types of products, both in early-phase clinical trials where the objective is to evaluate the activity of product using intermediate endpoints that can be observed in relatively short time, and in late-phase clinical trials where it may be difficult to demonstrate superiority with respect to overall survival compared to currently approved products due to switch-over of control to experimental treatment arm after disease progression.

4 Summary

Our current understanding of diseases at a molecular level, based in part on advances in genomics, has made it possible to further subdivide disease categories previously defined by site and histology, resulting in smaller populations to study new products. The current generation of products do not fit into the cytotoxic chemotherapy paradigm and require innovative thinking in designing, conducting and interpreting results from clinical trials. We must rethink the goal of each phase of clinical trials in the overall development of new drug products given the differing mechanisms of action and treatment effects of new targeted therapies. Future clinical trials are likely to be more complex and biomarker-based, with adaptive features. Simulation of such designs may become necessary to understand the operational complexities such that statistical properties such as type I error control and study power are not compromised. Further research is needed in identifying intermediate endpoints (for example, response criteria that would capture responses to immunotherapy) so that informative go-no-go decisions for further development of a product can be made.

Most of the clinical trials with time-to-event endpoints are designed assuming an exponential distribution of the outcome measure and a proportional hazard function. However it is not uncommon to observe that these assumptions are violated. Simulation of clinical trials where these assumptions do not hold may be useful in designing and planning such clinical trials. Ultimately, there should be a prospective statistical plan detailing alternative statistical methods for analyzing and summarizing the data should be in place if these assumptions do not hold true.

The selection of endpoints can become even more challenging if, for example, immunotherapy in combination with chemotherapy is being studied. A single intermediate endpoint in such circumstances may not capture the activity and effectiveness of both types of therapies. Careful consideration of the selection of endpoint, length of treatment and length of follow-up would be needed at the design stage.

The importance of timing and rigor in determining the analytic performance of the companion diagnostic test cannot be ignored with the advent of increasing number of targeted therapies. Understanding the statistical properties of the device such as sensitivity, specificity, positive and negative predictive values is essential.

Finally, with the limited number of patients and other resources, collaboration among pharmaceutical and device companies, academicians, government agencies including regulatory agencies, payers and patient advocacy groups is crucial in order to conduct future clinical trials that are informative as to the safe and effective use of a product. Statisticians are in a unique position to resolve the complexities inherent in the design of efficient and informative clinical trials.

References

1. Mile stones in U.S. Food and Drug law history. http://www.fda.gov/AboutFDA/WhatWeDo/History/Milestones/ucm128305.htm.
2. 21 Code of Federal Regulations, Part 314.126.
3. FDA Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products. 1998.
4. Simon R. Optimal two-stage designs for phase II clinical trials. Control Clin Trials. 1989;10:1–10.
5. Kane RC, et al. Sorafenib for the treatment of advanced renal cell carcinoma. Clin Cancer Res. 2006;12:7271–8.
6. Hodi FS, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010;363:711–23.
7. Chapman PB, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N Engl J Med. 2011;364:2507–16.
8. Hauschild A, et al. Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. Lancet. 2012;380:358–65.
9. Wakelee H, et al. Cooperative group research efforts in lung cancer 2008: focus on advanced-stage non-small-cell lung cancer. Clin Lung Cancer. 2008;9:346–51.
10. Freidlin B, Simon R. Adaptive signature design: an adaptive clinical trial design for generating and prospectively testing a gene expression signature for sensitive patients. Clin Cancer Res. 2005;11:7872–8.

11. Freidlin B, et al. The cross-validated adaptive signature design. Clin Cancer Res. 2010;16:691–8.
12. Redman MW, Crowley JJ, Herbst RS, Hirsch FR, Gandara DR. Design of a phase III clinical trial with prospective biomarker validation: SWOG S0819. Clin Cancer Res. 2012; 18(15):4004–12.
13. Simon N, Simon R. Adaptive enrichment designs for clinical trials. Biostatistics. 2013;14:613–25.
14. Mehta C, et al. Biomarker driven population enrichment for adaptive oncology trials with time to event endpoints. Stat Med. 2014;33:4515–31.
15. Barker AD, et al. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. Clin Phamacol Ther. 2009;86:97–100.
16. Herbst RS, et al. Lung master protocol (Lung MAP)—a biomarker-driven protocol for accelerating development of therapies for squamous cell lung cancer: SWOG S1400. Clin Cancer Res. 2015;21:1514–24.
17. National Cancer Institute Press Release. NCI-MATCH trial will link targeted cancer drugs to gene abnormalities. 2015. http://www.cancer.gov/news-events/press-releases/2015/nci-match.
18. Sridhara R, et al. Current statistical challenges in oncology clinical trials in the era of targeted therapy. Stat Biopharm Res. 2015;7(4):348–56. doi:10.1080/19466315.2015.1094673.
19. In vitro companion diagnostic devices: Guidance to Industry and Food and Drug Administration Staff. 2014. http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM262327.pdf.
20. Kazandjian D, et al. FDA approval summary: Nivolumab for the treatment of metastatic non-small cell lung cancer with progression on or after platinum-based chemotherapy. Onclologist. 2016;21:634–42.
21. Pembrolizumab product label: Section 14.1, Figure 2. http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/125514s004s006lbl.pdf.
22. Nivolumab product label: Section 14. http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/125554s001lbl.pdf.

http://www.springer.com/978-981-10-0124-6

Frontiers of Biostatistical Methods and Applications in Clinical Oncology

Matsui, S.; Crowley, J. (Eds.)
2017, XIII, 438 p. 89 illus., Hardcover
ISBN: 978-981-10-0124-6