Contemporary Clinical Trials 32 (2011) 446–452

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Analysis of tumor burden versus progression-free survival for Phase II decision making

Jane Fridlyand *, Lee D. Kaiser, Gwen Fyfe

Genentech, South San Francisco, CA, United States

ARTICLE INFO

Article history:
Received 7 October 2010
Accepted 17 January 2011

Keywords:
Phase II decision making
Tumor burden
Re-sampling
Oncology

ABSTRACT

Purpose: There have been recent recommendations to use percentage change in tumor burden (dTB) as a primary endpoint in randomized Phase II trials. We assessed whether dTB is better for the decision to start a Phase III trial than is progression-free survival (PFS).
Methods: We repeatedly sampled patients from six large randomized trials to obtain simulated Phase II trials. We derived PFS and dTB endpoints on the trial patients and determined the fraction of simulated trials with positive results for each endpoint. We supplemented these analyses with regression analyses to assess the ability of PFS and dTB to predict overall survival (OS).
Results: The best PFS endpoint included tumor assessments through 6 months after the last patient enrolled. With 70 patients in each simulated Phase II trial, the estimated rate of a correct 'Phase III go' decision ranged from 0.74 to 0.91 across the six parent studies. The best dTB endpoint was the last dTB through 6 months after the last patient enrolled, with corresponding rates of 0.54 to 0.81. The PFS rate was better than the dTB rate in five studies. PFS and dTB are individually statistically significant predictors of OS ($p<0.05$). In all six studies PFS added significantly to the regression models with dTB included, while in only two studies did dTB add significantly to the regression model with PFS included.
Conclusion: Analysis of PFS in randomized Phase II trials generally leads to better 'Phase III go' decisions than does analysis of dTB. Tumor burden analyses should be used in supportive analyses to a primary PFS analysis.

1. Introduction

Phase II is a critical stage in clinical drug development. Oncology drugs have a higher likelihood of progressing to Phase III than non-oncology products, but the success rate in Phase III for oncology drugs is notably lower [1]. Better analytical approaches for Phase II data that either improve the success rate or decrease the time to the 'Phase III go' decision could improve the efficiency of oncology drug development.

Recently, several authors have suggested that the observed or predicted percentage change in tumor burden (dTB) is predictive of overall survival, and thus can be used as an early and more comprehensive marker of drug response compared to traditional objective response or progression and thus serve as a primary endpoint in randomized Phase II trials [2–6]. Some specific quotations include the following:

- It is likely that the [randomized discontinuation trial] design can be made even more efficient with modifications, such as using the full range of changes in tumor burden after random assignment, rather than a binary arbitrary criterion of "progression" as the primary end point [2].
- Therefore, a scheduled visit at week 8 with computed tomography imaging for tumor size measurements can provide a critical signal for drug effect [4].

* Corresponding author at: 1 DNA Way, MS 441b, South San Francisco, CA 94080, United States. Tel.: +1 650 467 3755.
E-mail address: fridlyand.jane@gene.com (J. Fridlyand).

1551-7144/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.cct.2011.01.010

- Change in tumor size can be used as a primary end point in the design and evaluation of Phase II studies and in supporting go/no-go decisions and Phase III study design [6].

If this proposal is correct, the decision to study an agent in a Phase III trial could be made more quickly and potentially with fewer patients compared with a decision based on progression-free survival (PFS) in a traditional randomized Phase II trial.

Quantitative analysis presented in these published works primarily focuses on three main points: (1) Drug-independent models to predict the percentage change in tumor burden can be successfully constructed; (2) predicted dTB at a chosen landmark time is an independent predictor of survival; and (3) the landmark dTB is decreased with an active agent. However, none of the publications demonstrate the final claim that utilizing dTB as a primary endpoint in randomized Phase II trials leads to better decision making than does the use of PFS.

A good 'Phase III Go' rule would have a high probability of starting a Phase III trial with an active drug (the true positive rate) and conversely have a low probability of starting Phase III with an inactive drug (the false positive rate). We estimated these rates for dTB-based and PFS-based 'Phase III Go' rules through the sub-sampling of patients from six completed, large, randomized oncology trials to form simulated Phase II trials.

To explore reasons for our findings on Phase II tumor burden analyses, we also assessed the relationships of dTB and PFS with overall survival (OS) in order to quantify the relative contributions of dTB and PFS in predicting patients' overall survival.

2. Data and methods

2.1. Clinical trials

We used two Phase III studies and one large Phase II study of bevacizumab (3 studies), capecitabine, erlotinib, and trastuzumab (1 study each) in metastatic breast, colorectal and lung cancer indications (see Table 1). The studies were selected based on the availability of individual patient data and the desire to include different indications and treatments. The six studies spanned a diverse set of conditions to increase the assurance that the conclusions of the work are not driven by a specific indication, line of therapy, or mechanism of action of a drug. No study was explicitly excluded for any reason other than lack of accessible data for the analysis.

Radiological tumor assessments were performed on regular schedules with intervals between assessments of 6, 8, 9, or 12 weeks, depending on the study. Patient time of progressive disease was determined using RECIST criteria for the bevacizumab and erlotinib studies and using WHO criteria for the capecitabine trial and modified WHO for the trastuzumab trial. Tumor burden for our analysis was determined as the sum of longest diameters across the target lesions for the bevacizumab and erlotinib studies or across the marker lesions for the trastuzumab and capecitabine studies. All studies but AVF2119g were positive for PFS, and all but AVF2119g and AVF2192g were positive for OS. All studies were positive or nearly so for the overall response rate, which is the sum of confirmed partial and complete responses. AVF2119g was the only study that did not lead to the successful registration of the drug, and is considered to be a negative study for our analyses.

Patients were included in our analyses if they had a baseline tumor assessment, received at least one dose of study medication, and had at least one post-baseline tumor assessment. This differs from the usual intent-to-treat approach in the analysis of a Phase III study and is meant to mirror a more pragmatic approach taken in Phase II. Also, the treatment arm that contained 5-FU/leucovorin plus bevacizumab in AVF2107g was dropped from the study by design and is not included here. Thus, there were 2714 (82%) patients available for the analyses from the six studies, with numbers by study as shown in Table 1. Each study was treated separately throughout the manuscript, with no data pooled across these diverse trials.

2.2. Methods

We first confirmed that dTB, the percentage change from baseline TB, represents an appropriate adjustment for baseline for the comparison of treatment groups [13] through plotting of

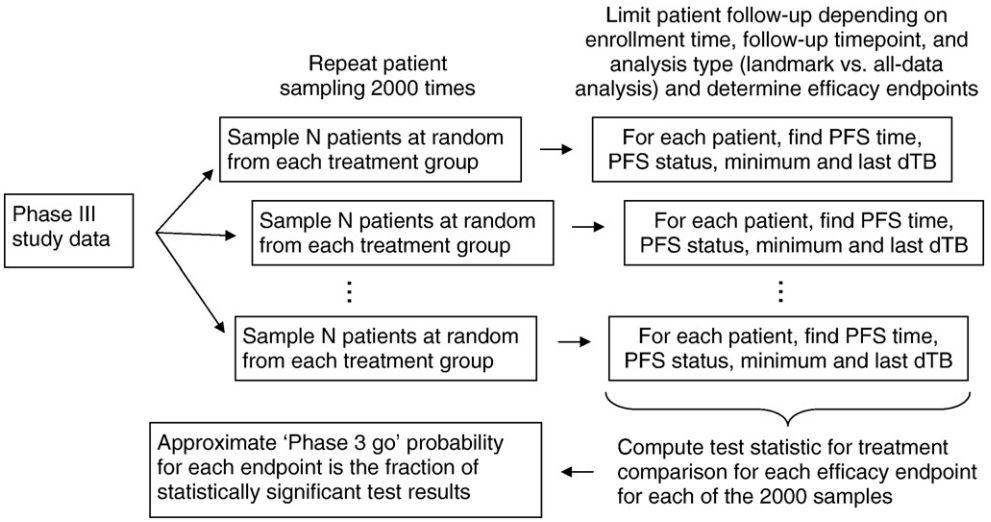

Fig. 1. Simulation of Phase II trials: patient selection and outcome determination. PFS, progression-free survival; dTB, percentage change from baseline tumor burden.