Using Change in Tumor Size as Primary Endpoint in Oncology Phase II studies

L. Claret (1), V. André (2), D. de Alwis (2), R. Bruno (1)

1 Pharsight Corporation  2 Eli Lilly & Co., Erl Wood, UK

Background

Small non-randomised Phase II trials have been the norm for the development of many Oncology compounds
  - These small trials have been associated with unreliable evidence to move to Phase III
  - Major criticism is that the interpretation relies on comparing results to an historical control

The use of Objective Response Rate (ORR) or Progression-Free Survival (PFS) as primary endpoint has been questioned [1]
  - Both ORR and PFS are based on RECIST criteria for tumor response evaluation, which consists in categorizing the change in tumor size observed
  - Using these endpoints, too many patients would be required to establish dose-response relationships or to compare alternative schedules
  - Trials using PFS often result in lengthy trials preventing rapid decision making

Change in tumor size (CTS) from baseline has been proposed to be used as the primary endpoint in Phase II studies [2, 3]
  - CTS as a continuous longitudinal endpoint can be considered as a biomarker for drug effect in early clinical studies
  - The use of this continuous patient-level endpoint rather than categorizing the changes is more sensitive in assessing treatment effect
  - Randomized studies to assess dose-response, optimal scheduling can therefore be envisaged

Time to Event Model for PFS

The distribution is shifted to the left to reach a 50% improvement in PFS

Log-Ratio for the Investigational Treatment

Lognormal distribution has the best fit

Observations: 173 Total; 17 Censored

Factors associated with an improvement in PFS:
  - Tumor Shrinkage
  - Decrease in baseline tumor size
  - ECOG Performance Status (0 vs. 1)

Log-Ratio of the Investigational Treatment

The distribution is shifted to the left to reach a 50% improvement in PFS

Note: Most of shifted CTS are below 0 (tumor shrinkage)

Conclusion

A time to event model was developed to predict PFS based on observed fractional change and patients characteristics.

Simulations showed that at least 60% tumor shrinkage should be achieved to observe 50% PFS improvement (a relevant clinical outcome).

Trial simulations of an investigational treatment versus docetaxel, with 120 patients using a 2:1 randomization, demonstrated that:
  - Using change in tumor size (log-ratio test) as an endpoint is more efficient than PFS
  - The power to detect a 2-month improvement in PFS is 60% with the log-rank test and 100% using the log-ratio test.
  - A 1 month improvement in PFS may be shown using the log-ratio test.

The model can be used to support interim futility analysis in a Phase II studies.

A disease-specific survival model can also be used to make inference on expected survival of the investigational treatment and to support go-no go decisions and Phase III study design [5, 6].

Using change in tumor size as a primary endpoint in Phase II oncology studies is an excellent alternative to the commonly used endpoints.

Methods

Data: Phase III of docetaxel vs. pemetrexed [4]
  - N=225 with ECOG 0 or 1 with at least one tumor measurement in docetaxel arm
  - 23 early dropout (10.2 %)
  - Delta CTS at 2 (median = 33 days) were used to build the model

Model: A parametric time to event model for PFS as a function of CTS (change in tumor size) and other prognostic factors was developed
  - Weibull, exponential, normal, lognormal, logistic, loglogistic distributions were tested to describe the distribution time to event (disease progression or death)
  - S-Plus Censor Reg (version 6.2)

Simulations: A randomized Phase II study of a new investigational treatment vs. docetaxel was simulated under various scenarios for the efficacy of the investigational treatment
  - PFS model was used to assess the CTS required to achieve the desired efficacy goals in term of PFS
  - Multiple replicates (n=1000) of study design (120 patients in a 2:1 randomisation) were simulated and study performance (% successful trials) was assessed to compare design and endpoints
  - A Log-rank test was used to compare PFS and t-test was used on the log ratio, assumed to be normally distributed (2)

Simulations Outcome

In all simulated scenarios, CTS was always more efficient than PFS (greater power with less patients).

References