Contemporary Clinical Trials 31 (2010) 620–623

Contents lists available at ScienceDirect

Contemporary Clinical Trials

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

Short communication

Assessing the success probability of a Phase III clinical trial based on Phase II data

Zheng Su *

Department of Biostatistics, Genentech Inc, Mail Stop 451b, 1 DNA Way, South San Francisco, CA 94080, USA

ARTICLE INFO

Article history:
Received 16 April 2010
Accepted 12 August 2010

Keywords:
Clinical trial
Phase II
Phase III
Bayesian
Likelihood

ABSTRACT

Assessing the probability that a Phase III clinical trial will demonstrate clinically relevant efficacy based on Phase II data is an important topic in clinical drug development. An accurate estimate of how likely a Phase III trial will succeed based on available data will inform the decision on whether to move an experimental medicine forward to Phase III testing. Bayesian and likelihood methodologies have been developed in the literature to assess the probability of reproducibility in clinical trials for parametric models. A class of approaches that combines the Bayesian and likelihood approaches is proposed to evaluate the success probability of a Phase III trial based on Phase II data, which applies to the parametric, semi-parametric, and non-parametric settings and includes the Bayesian and likelihood approaches as special cases.

1. Introduction

Assessing the probability that a Phase III clinical trial will demonstrate clinically relevant efficacy based on Phase II data is an important topic in clinical drug development. Usually a Phase II trial collects some preliminary data on the efficacy of a molecule, and whether it has clinically relevant efficacy will be rigorously tested in one or multiple Phase III trials if promising results are observed in the Phase II trial. An accurate estimate of the Phase III success probability will help understand the risk of further studying a molecule in a Phase III trial.

The probability that a second clinical trial with the same protocol will achieve statistical significance after the first trial produced a statistically significant result has been studied by Goodman [1] and Shao and Chow [2]. The objective of the second trial is to confirm that the significant result from the first trial is reproducible. Let H_0 be the null hypothesis that the molecule under investigation and the control have the same efficacy, and H_1 be the alternative hypothesis. For a one-sided alternative hypothesis, H_0 is rejected when $S>c$, where S is a test statistic and c is a constant. Discussions for two-sided hypotheses are similar. If H_1 has a parametric form, the power is given by

$$P(S > c|H_1) = P(S > c|\theta), \tag{1}$$

where θ is a parameter or vector of parameters. Goodman [1] and Shao and Chow [2] considered the reproducibility probability $P\left(S > c|\hat{\theta}\right)$, where $\hat{\theta}$ is a point estimate of the parameter θ (e.g. the maximum likelihood estimate (MLE) of θ) based on the previous trial. They also considered a Bayesian approach where they studied the reproducibility probability

$$P(S > c|\mathbf{x}) = \int P(S > c|\theta) f(\theta|\mathbf{x}) d\theta, \tag{2}$$

where $\mathbf{x}$ is the data set observed from the previous trial and $f(\theta|\mathbf{x})$ is the posterior density of θ given $\mathbf{x}$. Shao and Chow [2] also studied a confidence bound approach with the idea of finding a lower confidence bound $\hat{\theta}_{\_}$ of θ and evaluating a more conservative reproducibility probability $P\left(S > c|\hat{\theta}_{\_}\right)$.

The sample size for a Phase II trial is normally limited. As a result, the prior distribution in the Bayesian approach may heavily influence the posterior distribution, which could lead to a poor estimate of the Phase III success probability if the

* Tel.: +1 650 467 1190; fax: +1 650 467 4155.
E-mail address: su.zheng@gene.com.

1551-7144/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cct.2010.08.008

prior is poorly chosen. A clear disadvantage of the maximum likelihood approach is that the MLE may have a large variance and this point estimate does not reflect the sample size of the Phase II trial. These features of the Bayesian and likelihood approaches served as the motivations for the development of a class of approaches to assessing the Phase III success probability, which combines the Bayesian and likelihood approaches and include them as special cases.

2. A class of hybrid Bayesian-likelihood approaches

2.1. Parametric setting

In this section, we propose an approach that combines the Bayesian and likelihood approaches, where the degree of tradeoff between these two methods is controlled by a user-specified parameter $\rho\in(0,1)$. In particular, for given observed data $\mathbf{x}$ in the Phase II trial, $\rho\in(0,1)$, and an initial prior distribution $g(\theta)$ for the model parameter θ, we update the posterior distribution as follows:

$$\tilde{f}(\theta|\mathbf{x}) = \frac{g(\theta)L(\theta|\mathbf{x})I\left\{L(\theta|\mathbf{x})\geq\rho L\left(\hat{\theta}|\mathbf{x}\right)\right\}}{\int g(u)L(u|\mathbf{x})I\left\{L(u|\mathbf{x})\geq\rho L\left(\hat{\theta}|\mathbf{x}\right)\right\}du}, \tag{3}$$

where $\hat{\theta}$ is the MLE of θ based on $\mathbf{x}$. The inclusion of the additional term $I\left\{L(\theta|\mathbf{x})\geq\rho L\left(\hat{\theta}|\mathbf{x}\right)\right\}$ in the proposed approach addresses some of the difficulties of the Bayesian and likelihood approaches. Intuitively, since the maximum likelihood point estimate of θ calculated based on a small sample size has a large variance, the Bayesian methodology built in the new approach provides the smoothing through the prior distribution and thus reduces the variance. On the other hand, since a poorly chosen prior distribution may have large values on highly unlikely values of θ, by restricting the prior distribution on the set of θ's whose likelihood function values are greater than the threshold $\rho L\left(\hat{\theta}\right)$, only the part of the prior defined on promising values of θ will have an impact on the estimation of the posterior distribution. Note that in the two extreme cases where $\rho=0$ and $\rho=1$, this approach reduces to the standard Bayesian and likelihood approaches, respectively, and the defined density becomes a point mass when $\rho=1$. The Phase III success probability will then be estimated by

$$P(S\geq c|\mathbf{x}) = \int P(S\geq c|\theta)\tilde{f}(\theta|\mathbf{x})d\theta. \tag{4}$$

In cases where no prior knowledge of the parameter θ is available and we have a noninformative prior, this approach reduces to calculating the Phase III success probability for various possible true parameter values that correspond to sufficiently high likelihoods (i.e. higher than 100ρ percent of the maximum likelihood), and the final Phase III success probability is calculated based on an weighted average of these probabilities with the likelihood values being the weights.

The probability in (4) can be estimated without resorting to numerical integration. With a noninformative prior distribution the posterior distribution of θ can be estimated by $\theta|\mathbf{x}\sim N\left(\hat{\theta},\hat{\sigma}^2_{\hat{\theta}}\right)$, where $\hat{\theta}$ is the MLE based on $\mathbf{x}$ and $\hat{\sigma}^2_{\hat{\theta}}$ is the sample variance of $\hat{\theta}$. As a result, the distribution of $S(\theta)$ can be estimated by $S(\theta)|\mathbf{x}\sim N\left(S\left(\hat{\theta}\right),\hat{\sigma}^2_{S\left(\hat{\theta}\right)}\right)$, where $\hat{\sigma}^2_{S\left(\hat{\theta}\right)}$ denotes the sample variance of $S\left(\hat{\theta}\right)$ and can be approximated via the delta-method. The probability in (4) can be estimated by

$$P(S(\theta)\geq c) = P\left(Z>\frac{c-S\left(\hat{\theta}\right)}{\hat{\sigma}_{S\left(\hat{\theta}\right)}}\right) = 1-\Phi\left(\frac{c-S\left(\hat{\theta}\right)}{\hat{\sigma}_{S\left(\hat{\theta}\right)}}\right), \tag{5}$$

where Z denotes a standard normal random variable and Φ denotes the cumulative distribution function (cdf) of Z.

2.2. Semi-parametric setting

For clinical trials with failure time as the endpoint, the semi-parametric Cox's proportional hazards model has been widely used. Suppose we have m patients on the experimental treatment and n patients on the standard of care (SOC) with failure time as the endpoint. We assume the proportional hazards model $h_1(t)=e^{\beta}h_0(t)$ relating the hazard functions of the two treatments. For the survival outcome, denote the pooled data by $\left(\tilde{T}_i,\eta_i,Z_i\right), i=1,\ldots,m+n$, where $\tilde{T}_i=min(T_i,C_i)$, T_i is the actual survival time and C_i is the censoring time, $\eta_i=I_{\{T_i\leq C_i\}}$, and $Z_i=1$ or 0 for the experimental treatment or the SOC, respectively. The log partial likelihood of the hazard ratio β is

$$l(\beta) = \sum_{i=1}^{m+n}\eta_i\left\{\beta Z_i-log\left(\sum_{j:\tilde{T}_j\geq\tilde{T}_i}exp\left(\beta Z_j\right)\right)\right\}. \tag{6}$$

The hypotheses are $H_0:\beta=0$ versus $H_1:\beta<0$. The Wald statistic has the form $\hat{\beta}/\left(-\ddot{l}\left(\hat{\beta}\right)\right)^{\frac{1}{2}}$, where $\hat{\beta}$ maximizes the log partial likelihood $l(\beta)$ and $\ddot{l}$ denotes the second derivative of l. The null hypothesis H_0 can be evaluated based on the asymptotic normality of $\left(-\ddot{l}\left(\hat{\beta}\right)\right)^{-1/2}\left(\hat{\beta}-\beta\right)$ or the limiting χ^2 distribution of $2\left(l\left(\hat{\beta}\right)-l(\beta)\right)$. Under the Cox model, we will use the partial likelihood in place of the likelihood function in (3) to update the posterior distribution $f(\beta|\mathbf{x})$ for β, and the Phase III success probability will be evaluated as follows:

$$P\left(e^{\hat{\beta}_3}<c|\mathbf{x}\right) = \int P\left(e^{\hat{\beta}_3}<c|\beta\right)f(\beta|\mathbf{x})d\beta, \tag{7}$$

where $e^{\hat{\beta}_3}$ denotes the hazard ratio calculated from the Phase III data and c denotes the desired hazard ratio that is clinically relevant. To calculate $P\left(e^{\hat{\beta}_3}<c|\beta\right)$, Breslow's [3] estimator $\hat{\Lambda}_0(t)$ of the baseline cumulative hazard function $\Lambda_0(t)=\int_0^t h_0(s)ds$ based on the observed data $\left(\tilde{T}_i,\eta_i,Z_i\right), i=1,\ldots,m+n$ will be used, and the survival function for the SOC will be estimated via $\hat{S}_0(t)=exp\left(-\hat{\Lambda}_0(t)\right)$.

Similar to (5), the reproducibility probability in (7) can be estimated via

$$P(\beta<log(c)) = P\left(Z<\frac{log(c)-\hat{\beta}_2}{\left(-\ddot{l}\left(\hat{\beta}_2\right)\right)^{\frac{1}{2}}}\right) = \Phi\left(\frac{log(c)-\hat{\beta}_2}{\left(-\ddot{l}\left(\hat{\beta}_2\right)\right)^{\frac{1}{2}}}\right), \tag{8}$$

where $e^{\hat{\beta}_2}$ denotes the hazard ratio observed in the Phase II trial and $\beta \sim N\left(\hat{\beta}_2, -\ddot{l}\left(\hat{\beta}_2\right)\right)$.

Alternatively, without calculating $\ddot{l}\left(\hat{\beta}_2\right)$ the reproducibility probability in (7) can be estimated via

$$P(\beta < log(c)) = P\left(Z < \frac{log(c) - \hat{\beta}_2}{(4/N_2 + 4/N_3)^{\frac{1}{2}}}\right) = \Phi\left(\frac{log(c) - \hat{\beta}_2}{(4/N_2 + 4/N_3)^{\frac{1}{2}}}\right), \quad (9)$$

where N_2 denotes the number of events observed in the Phase II trial, N_3 denotes the number of events that the Phase III trial is powered with, and $\beta \sim N\left(\hat{\beta}_2, 4/N_2 + 4/N_3\right)$.

2.3. Nonparametric setting

When the distribution under the alternative hypothesis does not have a parametric form, neither the Bayesian nor the proposed hybrid Bayesian-likelihood approach applies since there is no parameter to estimate in this case. To circumvent this difficulty, the nonparametric empirical likelihood proposed by Owen [4–6] can be applied in place of the likelihood or partial likelihood function in the non-parametric setting.

We consider the one-sample case with a single cdf. For the underlying cdf F of $\mathbf{x}$, where $\mathbf{x} = (x_1,\ldots,x_n)$ denotes the data collected from the Phase II trial, the nonparametric likelihood of F is defined as

$$L(F) = \prod_{i=1}^{n} (F(x_i) - F(x_i -))$$

where $F(x) = P(X \le x)$ and $F(x-) = P(X < x)$. It is easy to show that the nonparametric likelihood is maximized by the empirical cumulative distribution function (ecdf) $F_n(x)$. Thus, the ecdf is the non-parametric maximum likelihood estimate (NPMLE) of F. To have a positive nonparametric likelihood, a distribution must place a positive probability on every one of the observed data values. We may consider the Phase III success probability $P\left(S > c \mid \tilde{F}\right)$ for any distribution eF that places probability p_i on x_i, $\sum_{i=1}^{n} p_i = 1$ with the constraint $\prod_{i=1}^{n} p_i > \rho L\left(\hat{F}_n\right) = \rho(1/n)^n$, and the non-parametric problem is essentially reduced to a parametric problem with the parameter vector $\theta = (p_1,\ldots,p_n)$. When n is large estimating the corresponding n-dimensional integral is computationally intensive, and various Monte Carlo methods can be applied in this setting to reduce computational cost; see Liu [7]. Note that the same idea applies to two-sample problems.

3. Example

In this section, we consider a two-sample problem with binary endpoints. Suppose a Phase II trial has been completed comparing an experimental drug with the SOC, and promising results have been observed with n patients treated with each drug inthis trial. Denote the observed response rates for the experimental drug and the SOC in the Phase II trial by $\hat{p}_1$ and $\hat{p}_2$, respectively, and $\hat{p}_1 - \hat{p}_2 = 20\%$ with $\hat{p}_1 = 40\%$ and $\hat{p}_2 = 20\%$. A response rate difference of 15% or more favoring the experimental drug is considered clinically meaningful and is desired to be demonstrated in the Phase III trial, and it is important to assess, based on the Phase II results, the probability of observing a difference of this magnitude in the Phase III trial. The Phase III trial has been sufficiently powered with a sample size of 750 patients per arm. We used Eqs. (4) and (5) to estimate the success probability of the Phase III trial for 4 difference scenarios with various sample sizes ($n = 40,60,80,100$). Intuitively, the success probability of the Phase III trial should increase as the sample size of the Phase II trial increases since a larger Phase II sample size means that the observed promising results in the Phase II trial are less likely due to chance. For Eq. (4) $\rho = 0.1$ was used with a non-informative prior distribution. In this example $\hat{\theta} = (\hat{p}_1, \hat{p}_2)$, $\theta = (p_1,p_2)$, where $p_i, i = 1,2$ are the respective unknown response rates for the experimental drug and the SOC, and $S(\theta) = p_1 - p_2$.

We can see from Table 1 that the probability of Phase III success increases as the sample size of the Phase II trial increases, and Eqs. (4) and (5) gave probabilities that differ by only 1% for all the cases considered. The Phase III success probability estimated via the Bayesian approach is relatively smaller and the likelihood approach produced the same success probability for all the cases considered as this approach does not take into account the sample size of the Phase II trial.

4. Conclusions

Assessing the probability that a Phase III clinical trial will demonstrate clinically relevant efficacy based on Phase II data is an important topic in clinical drug development. An accurate estimate of how likely a Phase III trial will succeed base on available data will inform the decision on whether to move an experimental medicine forward to Phase III testing. In this paper, a class of approaches that combines the Bayesian and likelihood approaches has been proposed to evaluate the success probability of a Phase III trial based on Phase II data, which applies to the parametric, semi-parametric, and non-parametric settings and includes the Bayesian and likelihood approaches as special cases. Without resorting to numerical integration a simple estimate of the success probability of a Phase III trial based on asymptotic normality has also been developed. Simulation results show that the estimate performed similarly to the proposed hybrid Bayesian-likelihood approach.

Acknowledgements

The authors thank the two anonymous reviewers for their insightful comments and suggestions, which have led to an improved paper.

Table 1
Probability of observing a ≥15% response rate increment for the experimental drug over the SOC in the Phase III trial based on Phase II data.

n	Eq. (4) (%)	Eq. (5) (%)	Bayesian (%)	Likelihood (%)
40	68	69	66	98
60	72	73	70	98
80	75	76	73	98
100	78	79	76	98

References

[1] Goodman SN. A comment on replication, *p*-values and evidence. Statistics in Medicine 1992;11:875–9.
[2] Shao J, Chow SC. Reproducibility probability in clinical trials. Statistics in Medicine 2002;21:1727–42.
[3] Breslow N. Covariance analysis of cencored surrival data. Biometrics 1974;30:89–99.
[4] Owen AB. Empirical likelihood ratio confidence intervals for a single functional. Biometrika 1988;75:237–49.
[5] Owen AB. Empirical likelihood ratio confidence regions. Annals of Statistics 1990;18:90–120.
[6] Owen AB. Empirical likelihood. New York: Chapman & Hall/CRC; 2001.
[7] Liu J. Monte Carlo strategies in scientific computing. New York: Springer; 2008.