Contemporary Clinical Trials Communications 12 (2018) 90–91

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

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

# Reconsideration of sample size and power calculation for overall survival in cancer clinical trials

Inkyung Jung[a], Hee Jung Ko[a], Sun Young Rha[b], Chung Mo Nam[c,*]

[a] Division of Biostatistics, Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, South Korea
[b] Division of Medical Oncology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul, South Korea
[c] Department of Preventive Medicine, Yonsei University College of Medicine, Seoul, South Korea

ARTICLE INFO

*Keywords:*
Progression-free survival (PFS)
Survival post-progression (SPP)
Hypo-exponential distribution

ABSTRACT

When designing a cancer clinical trial, it is usual to assume an exponential distribution for a time-to-event outcome such as overall survival (OS). OS is often expressed as the sum of progression-free survival (PFS) and survival post-progression (SPP), each of which is assumed to be exponentially distributed. Then, OS does not follow an exponential distribution any more but a gamma or hypo-exponential distribution. In this study, we derived a sample size calculation formula for comparing OS between two treatment arms using the log-rank test for OS following a gamma or hypo-exponential distribution. We conducted a simulation study to evaluate the sample size and power calculation based on the gamma or hypo-exponential distribution. We found that we could reduce the sample sizes considerably compared to when assuming an exponential distribution for OS.

In cancer clinical trials, whether a benefit in progression-free survival (PFS) translates into overall survival (OS) is controversial. OS is the gold standard for clinical benefits in oncology [1], but PFS is often used as a primary endpoint because it is quicker to assess and is expected to correlate with OS. Giessen et al. [2] reviewed 50 randomized first-line trials in metastatic cancer and concluded that PFS consistently highly correlated with OS. However, it is often observed that a significant improvement in PFS fails to lead to a significant improvement in OS. Booth and Eisenhauer [3] explained that the most widely cited reason is the influence of post-progression therapy. Broglio and Berry [4] conducted a simulation study to examine the impact of survival post-progression (SPP) on OS by partitioning OS into PFS and SPP, i.e., OS = PFS + SPP. They assumed a treatment benefit in PFS and no treatment effect on SPP. Their simulation results showed that longer SPP resulted in a weaker correlation between the hazard ratios for PFS and for OS and a less significant OS hazard ratio.

On the contrary, as noted in the paper by Morita et al. [5], statistically significant efficacy in OS was found but not in PFS in certain clinical trials [6,7]. Morita et al. [5] conducted similar simulations to Broglio and Berry [4] under the different assumption that only SPP, and not PFS, differed between treatment arms. Their simulation results suggested that shorter PFS resulted in more statistically significant OS benefit, which means that the OS curves depended greatly on the difference in SPP.

In both simulation studies mentioned above, OS was expressed as the sum of PFS and SPP, each of which was assumed to be exponentially distributed. In most clinical trials where OS is the primary endpoint, sample size and power calculation is usually based on the assumption of exponentially distributed OS. However, is it reasonable to assume an exponential distribution for OS when it is the sum of two exponentially distributed variables? In this study, we discuss issues on the sample size and power calculation by properly considering the distribution of OS as the sum of PFS and SPP.

Assume that PFS and SPP are independently exponentially distributed with hazard rate $\lambda_1$ and $\lambda_2$, respectively. Defining OS as the sum of two independent variables following exponential distributions makes the distribution of OS a gamma distribution if $\lambda_1 = \lambda_2 = \lambda$, and hypo-exponential otherwise. The corresponding hazard function is given by

$$h_{OS}(t) = \begin{cases} \frac{\lambda^2 t}{1+\lambda t} & \text{if } \lambda_1 = \lambda_2 = \lambda \\ \frac{\lambda_1\lambda_2(e^{-\lambda_1 t} - e^{-\lambda_2 t})}{\lambda_2 e^{-\lambda_1 t} - \lambda_1 e^{-\lambda_2 t}} & \text{if } \lambda_1 \neq \lambda_2 \end{cases}. \quad (1)$$

It is obvious that the hazard functions of OS for two treatment arms are not proportional from equation (1).

Now we derive a sample size formula when OS is assumed to follow a gamma or hypo-exponential distribution. Let $h_{OS}^k(t)$, $f^k(t)$, $F^k(t)$, *and* $S^k(t)$ be the hazard function, probability density function, distribution function, and survival function of OS, and let $H^k(t)$ be the distribution function of censoring time for treatment

* Corresponding author.
*E-mail address:* cmnam@yuhs.ac (C.M. Nam).

https://doi.org/10.1016/j.conctc.2018.09.007
Received 19 June 2018; Received in revised form 12 September 2018; Accepted 24 September 2018
Available online 28 September 2018

**Table 1**
Estimated sample size for comparing overall survival (OS) between two treatment arms using log-rank test for OS following a gamma or hypo-exponential distribution ($\alpha = 0.05$, $\beta = 0.2$).

| Maximum follow-up time (mo) | Median PFS (mo) | | Median SPP (mo) | Gamma or Hypo-exponential | | Exponential | Sample size ratio (C)/(D) |
|---|---|---|---|---|---|---|---|
| | Arm A | Arm B | | $n_A = n_B$ (C) | Power (%) | $n_A = n_B$ (D) | |
| 36 | 9 | 6 | 3 | 122 | 81.7 | 222 | 0.55 |
| 48 | | | 6 | 171 | 81.1 | 358 | 0.48 |
| 60 | | | 9 | 254 | 80.6 | 529 | 0.48 |
| 60 | | | 12 | 367 | 80.4 | 776 | 0.47 |
| 36 | 9 | 3 | 3 | 24 | 86.9 | 38 | 0.63 |
| 48 | | | 6 | 37 | 83.9 | 68 | 0.54 |
| 60 | | | 9 | 58 | 81.9 | 106 | 0.55 |
| 60 | | | 12 | 85 | 79.8 | 160 | 0.53 |

arm $k$ ($k$ = A, B). Also denote the proportion of patients assigned to arm $k$ by $P^k$. Here we assume the equal number of patients for the two arms ($P^A = P^B = 0.5$) and the identical distribution function for censoring time ($H^A(t) = H^B(t) = H(t)$).

Schoenfeld [8] showed that the distribution of the log-rank test statistic is asymptotically normal with unit variance and mean given by

$$\varnothing = \frac{\sqrt{n}\int_0^\infty log\left\{\frac{h^A_{OS}(t)}{h^B_{OS}(t)}\right\}\pi(t)\{1-\pi(t)\}V(t)dt}{\sqrt{\int_0^\infty \pi(t)\{1-\pi(t)\}V(t)dt}}$$

where

$$\begin{aligned} V(t) &= P^B f^B(t)\{1-H^B(t)\} + P^A f^A(t)\{1-H^A(t)\} \\ &= 0.5\{1-H(t)\}\{f^A(t)+f^B(t)\} \end{aligned}$$

and

$$\begin{aligned} \pi(t) &= \frac{P^A\{1-F^A(t)\}\{1-H^A(t)\}}{P^A\{1-F^A(t)\}\{1-H^A(t)\} + P^B\{1-F^B(t)\}\{1-H^B(t)\}} \\ &= \frac{S^A(t)}{S^A(t)+S^B(t)}. \end{aligned}$$

Therefore, the total sample size $n$ for the log-rank test with significance level $\alpha$ and power $1-\beta$ can be derived as

$$n = \frac{(z_{1-\alpha/2} + z_{1-\beta})^2}{\left[\frac{\int_0^\infty log\left\{\frac{h^A_{OS}(t)}{h^B_{OS}(t)}\right\}\pi(t)\{1-\pi(t)\}V(t)dt}{\sqrt{\int_0^\infty \pi(t)\{1-\pi(t)\}V(t)dt}}\right]^2} \quad (2)$$

where $z_\gamma$ denotes the $\gamma \times 100$ percentile of the standard normal distribution. If the distributions of PFS and SPP are exponential, we can obtain explicit forms for all functions except for $H(t)$ in equation (2). We assume that the patient accrual follows a discrete uniform distribution $(0,\ \tau_0]$, and that censoring occurs only at the end of follow-up time $\tau$. Then, the integrand in the denominator can be expressed as a function of $t$ and the sample size $n$ can be calculated using Monte Carlo integration.

To evaluate the sample size and power based on formula (2), we performed a simulation study. Using the "hit-or-miss" method [9] for Monte Carlo integration, we calculated sample sizes for detecting a 3 or 6 months difference in medians of OS between two arms with $\alpha = 0.05$ and $1-\beta = 0.8$. We compared the results with the sample sizes calculated assuming exponentially distributed OS. Next, we generated OS using the sum of PFS and SPP for as many patients as the sample sizes calculated as described above. We generated each patient's PFS from an exponential distribution with a median of 9 for arm A and with a median of 6 or 3 months for arm B, and each patient's SPP from an exponential distribution with a median of 3, 6, 9, and 12 months for both arms. We assumed $\tau_0$ equal to 12 months and $\tau$ about 3 times of PFS + SPP for arm A. We estimated power of the log-rank test for OS using the simulated data from 1000 replications.

Table 1 summarizes the sample size calculation and power simulation results. Sample sizes calculated from equation (2) were much smaller than those calculated when assuming exponentially distributed OS. Estimated power of the log-rank test for OS following a gamma or hypo-exponential distribution was close to assumed power of 80% or a bit higher. The sample size calculated assuming exponentially distributed OS with a median of the sum of medians of exponentially distributed PFS and SPP is overestimated. This is due to a larger variance of an exponential distribution compared to a gamma or hypo-exponential distribution with the same mean. Our study results imply that we can reduce the sample size in cancer clinical trials with OS as the primary endpoint, by properly deriving the distribution of OS from the distributions of PFS and SPP.

We only considered that OS was expressed as the sum of independently exponentially distributed PFS and SPP. This may seem unreasonable. However, the study by Sundar et al. [10] that evaluated the relationship between SPP and PFS in advanced ovarian cancer showed that increases in median PFS generally lead to little change in SPP. Although little correlation between PFS and SPP was found in trials of a specific cancer, we do not believe that assuming independence between PFS and SPP is too unrealistic. Still, it would be worth trying to extend our results to more general situations. Considering that a patient can die without the occurrence of progression, further studies may be needed to derive a mixture distribution or to consider a multi-state model.

## References


[1] A. Chakravarty, R. Sridhara, Use of progression-free survival as a surrogate marker in oncology trials: some regulatory issues, Stat. Methods Med. Res. 17 (5) (2008) 515–518.
[2] C. Giessen, R.P. Laubender, D.P. Ankerst, et al., Progression-free survival as a surrogate endpoint for median overall survival in metastatic colorectal cancer: literature-based analysis from 50 randomized first-line trials, Clin. Canc. Res. 19 (1) (2013) 225–235.
[3] C.M. Booth, E.A. Eisenhauer, Progression-free survival: meaningful or simply measurable? J. Clin. Oncol. 30 (10) (2012) 1030–1033.
[4] K.R. Broglio, D.A. Berry, Detecting an overall survival benefit that is derived from progression-free survival, J. Natl. Cancer Inst. (Bethesda) 101 (23) (2009) 1642–1649.
[5] S. Morita, K. Sakamaki, G. Yin, Detecting overall survival benefit derived from survival postprogression rather than progression-free survival, J. Natl. Cancer Inst. (Bethesda) 101 (23) (2009) 1642–1649.
[6] J. Cortes, J. O'Shaughnessy, D. Loesch, et al., Eribulin monotherapy versus treatment of physician's choice in patients with metastatic breast cancer (EMBRACE): a phase 3 open-label randomised study, Lancet 377 (9769) (2011) 914–923.
[7] V. Heinemann, L.F. von Weikersthal, T. Decker, et al., FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial, Lancet Oncol. 15 (10) (2014) 1065–1075.
[8] D. Schoenfeld, The asymptotic properties of nonparametric tests for comparing survival distributions, Biometrika 68 (1) (1981) 316–319.
[9] O. Jones, R. Maillardet, A. Robinson, Introduction to Scientific Programming and Simulation Using R, CRC Press, Boca Raton, 2009.
[10] S. Sundar, J. Wu, K. Hillaby, J. Yap, R. Lilford, A systematic review evaluating the relationship between progression free survival and post progression survival in advanced ovarian cancer, Gynecol. Oncol. 125 (2) (2012) 493–499.