PHARMACEUTICAL STATISTICS
Pharmaceut. Statist. **9**:100–112 (2010)
Published online 4 June 2009 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/pst.380

Sample size considerations for Japanese patients in a multi-regional trial based on MHLW guidance

MAIN PAPER

Hui Quan*,†, Peng-Liang Zhao, Ji Zhang, Martin Roessner and Kyo Aizawa
Biostatistics and Programming, Sanofi-Aventis, Bridgewater, NJ, USA

Since the publication of the International Conference on Harmonization E5 guideline, new drug approvals in Japan based on the bridging strategy have been increasing. To further streamline and expedite new drug development in Japan, the Ministry of Health, Labour and Welfare, the Japanese regulatory authority, recently issued the 'Basic Principles on Global Clinical Trials' guidance to promote Japan's participation in multi-regional trials. The guidance, in a Q&A format, provides two methods as examples for recommending the number of Japanese patients in a multi-regional trial. Method 1 in the guidance is the focus of this paper. We derive formulas for the sample size calculations for normal, binary and survival endpoints. Computations and simulation results are provided to compare different approaches. Trial examples are used to illustrate the applications of the approaches. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: ***bridging study; ethnic effect; delta method; survival endpoint; risk reduction***

1. INTRODUCTION

Differences in ethnicity, culture and clinical practice between regions may have impact on efficacy, safety and dosing regimen of a drug, although such differences are mostly quantitative in nature. Nevertheless, duplications of large clinical studies in all regions will not only demand valuable resources but also delay the availability of new treatments to patients in the regions. To facilitate the development and registration of new medicines across regions, International Conference on Harmonization (ICH) E5 guideline [1] was issued in 1998 recommending a framework for evaluating the impact of ethnic factors on drug effects. In general, data developed in one region could be possibly extrapolated to the new region if there is a bridging study that shows evidence that the drug will behave similarly across regions.

The application of bridging to therapeutic products can be complicated. There are no

*Correspondence to: Hui Quan, Biostatistics and Programming, Sanofi-Aventis, BX2-403A, 200 Crossing Boulevard, PO Box 6890, Bridgewater, NJ 08807, USA.
†E-mail: hui.quan@sanofi-aventis.com

predefined acceptable statistical criteria in the ICH E5 for a bridging study to evaluate the similarity of efficacy in different populations. Shih [2] considered an approach for assessing consistency by ensuring the predictive probability of data from the new region falling within the previous experience. Chow *et al*. [3] proposed the use of a sensitivity index as a possible criterion for determining whether a bridging study is necessary, and a statistical method for the assessment of similarity of clinical results between regions using the concept of population bioequivalence. Hsiao *et al*. [4] used a group sequential analysis technique to evaluate the internal validity by assuming sequential availability of data from the original region and the new region.

Based on Uyama *et al*. [5], since the publication of ICH E5 in 1998, new drug approvals in Japan based on the bridging strategy have been gradually increasing, from 3.2% in 1999 to 25% in 2003. However, since most bridging studies are conducted after the new drugs' approvals in the original region, new drug approvals in Japan are lagging behind other countries by a number of years. Japanese patients cannot access effective and safe medicines at the same time as patients in other counties. To streamline and expedite new drug development in Japan, the Ministry of Health, Labour and Welfare (MHLW), the Japanese regulatory authority, issued the 'Basic Principles on Global Clinical Trials' guidance [6] in 2007 to promote Japan's participation in multi-regional trials. With the inclusion of a sufficient number of Japanese patients in these trials, it is then possible to assess potential ethnic differences within the trials.

The guidance outlines the basic concepts for planning and implementing the multi-regional trials in a Q&A format with one question specifically for considering the number of Japanese patients. Even though it does not particularly recommend any methods for deciding on the number of Japanese subjects in a multi-regional trial for establishing the consistency of treatment effects for the entire group and the Japanese group, it does provide two methods as examples. Kawai *et al*. [7] discussed sample size consideration based on Method 2. Method 1, which is the focus of this paper, can be described as follows.

Suppose that the treatment effect from a multi-regional trial for the entire group is D_{all} and for Japanese patients is D_{J}. Then the sample size for Japanese patients in the trial should satisfy

$$\Pr(D_{\text{J}}/D_{\text{all}} > \pi) \geqslant 1 - \beta' \tag{1}$$

where π is 0.5 or greater and $1-\beta'$ is 0.8 or greater. That is, the sample size for Japanese patients has to ensure $1-\beta'$ probability to demonstrate that the treatment effect for Japanese patients is more than a fraction of π of the overall treatment effect. Sample size calculations are conducted at the design stage. If at the design stage, there is a strong belief that treatment effects for Japanese patients and the other patients will be substantially different for the disease and the treatment under investigation, (1) should not be used and probably a separate trial should be conducted for Japanese patients. Therefore, we will assume similar or only slightly different treatment effects between Japanese patients and the other patients at the design stage for the above sample size derivations in this paper.

Since the guidance was issued recently, there is not much published literature discussing the experiences of using this method. Uesaka [8,9] first discussed the topic before issuing the guidance. Using simulation, Sekiguchi *et al*. [10] presented the sample size determination regarding the number of Japanese patients for a multi-regional oncology trial. In this paper, we provide a systematic and comprehensive discussion on how to estimate the sample size of Japanese patients for a multi-regional trial. Formulas are derived for normal, binary and survival endpoints. Section 2 provides the case of a normal endpoint with two approaches. The required sample size based on the simpler naïve approach is much greater than that of the other approach. With slight modifications, formulas for the normal endpoint can be applied to a binary endpoint. Section 3 presents the methodology for a survival endpoint. Two delta methods with three approaches for each delta method are considered. Section 4 presents the

DOI: 10.1002/pst

methodology of transforming the number of events for a survival endpoint to sample size. After the sample sizes are obtained, Section 5 presents the simulation results for comparing the performances of the two delta methods. Examples are used to illustrate the applications of the approaches in Section 6.

2. NORMAL ENDPOINT

Suppose that N_J and N_{NJ} are the sample sizes per treatment group, δ_J and δ_{NJ} are the true treatment effects and $\hat{\delta}_J$ and $\hat{\delta}_{NJ}$ are the corresponding estimates for Japanese and non-Japanese patients, respectively; $N = N_J + N_{NJ}$ is the overall sample size per treatment group determined to ensure sufficient overall power; $\hat{\delta}_{all} = (N_J\hat{\delta}_J + N_{NJ}\hat{\delta}_{NJ})/N$ is the estimate of the treatment effect for the entire group; the study endpoint follows a normal distribution with σ as the standard deviation. If the overall sample size N is for a two-sided test with significance level α and power $1-\beta$ for detecting an overall difference of δ, then

$$N = \frac{2(z_{1-\alpha/2} + z_{1-\beta})^2\sigma^2}{\delta^2} \tag{2}$$

where $z_{1-\varepsilon}$ is the $(1-\varepsilon)\times 100$ percentile of the standard normal distribution. Assuming positive values implying a better study outcome, δ in (2) should be positive. Since probability

$$\begin{aligned}&\Pr(\hat{\delta}_{all} \leqslant 0 | \delta_{all} = \delta)\\ &= \Pr\left(\sqrt{\frac{N}{2}}\frac{\hat{\delta}_{all} - \delta}{\sigma} \leqslant -\sqrt{\frac{N}{2}}\frac{\delta}{\sigma}\middle|\delta_{all} = \delta\right)\\ &= \Pr\left(Z \leqslant -\sqrt{\frac{N}{2}}\frac{\delta}{\sigma}\right) = \Phi(-z_{1-\alpha/2} - z_{1-\beta})\end{aligned}$$

is very close to zero (it is 0.0006 if $\alpha = 0.05$ and $1-\beta = 0.90$), where Z and Φ are the standard normal random variable and the cumulative distribution, respectively, (1) is essentially

$$\begin{aligned}&\Pr(\hat{\delta}_J/\hat{\delta}_{all} > \pi | \delta_J, \delta_{NJ})\\ &= \Pr(\hat{\delta}_J > \pi\hat{\delta}_{all} | \delta_J, \delta_{NJ}) \geqslant 1 - \beta'\end{aligned} \tag{3}$$

With

$$\hat{\theta} = (N - \pi N_J)\hat{\delta}_J - \pi N_{NJ}\hat{\delta}_{NJ} \sim N(\theta, \omega^2)$$

where $\theta = (N - \pi N_J)\delta_J - \pi N_{NJ}\delta_{NJ}$ and

$$\begin{aligned}\varpi^2 &= \sigma^2\left(\frac{2(N - \pi N_J)^2}{N_J} + 2\pi^2 N_{NJ}\right)\\ &= 2\sigma^2 N\frac{N + (\pi^2 - 2\pi)N_J}{N_J}\end{aligned} \tag{4}$$

(3) becomes

$$\begin{aligned}\Pr(\hat{\theta} > 0 | \delta_J, \delta_{NJ}) = \Pr((\hat{\theta} - \theta)/\varpi > \\ - \theta/\varpi | \delta_J, \delta_{NJ}) \geqslant 1 - \beta'\end{aligned}$$

or

$$\frac{(N - \pi N_J)\delta_J - \pi N_{NJ}\delta_{NJ}}{\varpi} \geqslant z_{1-\beta'} \tag{5}$$

Suppose that $\delta_J = u\delta_{NJ}$, with u around 1. The true effects for Japanese patients and the other patients are the same if $u = 1$. Let f_u ($N_{uJ} = f_u N$) be the corresponding minimum fraction of the Japanese patients, which satisfies (5). Then $\delta = f_u\delta_J + (1 - f_u)\delta_{NJ} = (uf_u + 1 - f_u)\delta_{NJ}$, $N_{uJ} = f_u N$ and $N_{uNJ} = (1 - f_u)N$. From (2),

$$N = \frac{2(z_{1-\alpha/2} + z_{1-\beta})^2\sigma^2}{(1 + (u-1)f_u)^2\delta_{NJ}^2}$$

From (4) and (5), f_u satisfies

$$\frac{(z_{1-\alpha/2} + z_{1-\beta})\sqrt{f_u}(u - \pi - \pi(u-1)f_u)}{(1 + (u-1)f_u)\sqrt{1 + (\pi^2 - 2\pi)f_u}} = z_{1-\beta'} \tag{6}$$

Since N of (2) is used for deriving f_u, f_u is not just based on (1) alone. It has a strong connection with the overall sample size for having sufficient power to detect a significant overall treatment effect. For a general u, there is no closed-form solution for f_u via (6). For given α, β, β', π and u, a numerical solution can be derived without much difficulty. If $u = 1$, i.e. $\delta_J = \delta_{NJ} = \delta$, a closed-form solution

DOI: 10.1002/pst

can be obtained as follows:

$$f_1 = \frac{z_{1-\beta'}^2}{(z_{1-\alpha/2} + z_{1-\beta})^2(1-\pi)^2 + z_{1-\beta'}^2(2\pi - \pi^2)} \quad (7)$$

To simplify the calculation, treating $\hat{\delta}_{\text{all}}$ as the fixed δ in (3), some practitioners may interpret (3) as

$$\Pr(\hat{\delta}_{\text{J}} > \pi\delta | \delta_{\text{J}} = \delta_{\text{NJ}} = \delta_{\text{all}} = \delta) \geqslant 1 - \beta' \quad (8)$$

Then the sample size for Japanese patients would be

$$N_{\text{J}} \geqslant \frac{2\sigma^2 z_{1-\beta'}^2}{(1-\pi)^2\delta^2} \quad (9)$$

which is greater than $f_1 N = N_{1\text{J}}$ as long as $\pi < 2$. This is unexpected since one would hope to reduce the required Japanese sample size if $\hat{\delta}_{\text{all}}$ is set to the true value δ. As demonstrated by the results in Table I, this approach could be very conservative and give sample size much bigger than necessary. Therefore, we strongly recommend practitioners to avoid the approach. For N of (2), (9) becomes

$$\begin{aligned} N_{\text{J}} &\geqslant \frac{z_{1-\beta'}^2 N}{(z_{1-\alpha/2} + z_{1-\beta})^2(1-\pi)^2} = N'_{1\text{J}} = f'_1 N \\ &= \left(1 + \frac{(2\pi - \pi^2) z_{1-\beta'}^2}{(z_{1-\alpha/2} + z_{1-\beta})^2(1-\pi)^2}\right) f_1 N \geqslant N_{1\text{J}} \end{aligned} \quad (10)$$

Table I presents the values of $f_{0.9}$, f_1, $f_{1.1}$ and f'_1 in (6), (7) and (10) for $\alpha = 0.05$ and various π, $1-\beta$ and $1-\beta'$. Naturally, $f_{0.9}$ is always greater than f_1, while $f_{1.1}$ is always smaller than f_1. When $\pi = 0.5$, $1-\beta = 0.9$ and $1-\beta' = 0.8$, $f_1 = 0.224$ or $N_{1\text{J}} = 0.224N$. That is, the Japanese sample size has to be at least 22.4% of the overall sample size when the overall sample size is chosen to have 90% power. The corresponding $f'_1 = 0.270$ or $N'_{1\text{J}} = 0.270N$ is 20.6% ((0.270–0.224)/0.224) higher than what is required by (3). Results in Table I also show that f_1 decreases as the overall power $1-\beta$ for the study increases. However, since the overall sample size also increases with the power, the actual sample size for Japanese patients may not be reduced. It is interesting to see that, when $\pi = 0.7$ and $1-\beta' = 0.9$, f'_1 has to be greater than 1 in order for

Table I. Values of $f_{0.9}$, f_1, $f_{1.1}$, f'_1, ρ and Ψ ($\alpha = 0.05$).

π	$1-\beta$	$1-\beta'$	$f_{0.9}$	f_1	$f_{1.1}$	f'_1	$(1-\beta)(1-\beta')$	ρ	Ψ
0.5	0.90	0.80	0.290	0.224	0.174	0.270	0.720	0.260	0.735
0.5	0.95	0.80	0.248	0.187	0.143	0.218	0.760	0.233	0.768
0.5	0.90	0.85	0.383	0.313	0.253	0.409	0.765	0.320	0.781
0.5	0.95	0.85	0.334	0.265	0.209	0.331	0.808	0.288	0.816
0.5	0.90	0.90	0.494	0.426	0.361	0.625	0.810	0.395	0.826
0.5	0.95	0.90	0.437	0.367	0.303	0.506	0.855	0.356	0.864
0.6	0.90	0.80	0.396	0.311	0.240	0.421	0.720	0.260	0.735
0.6	0.95	0.80	0.349	0.265	0.198	0.341	0.760	0.233	0.768
0.6	0.90	0.85	0.496	0.416	0.340	0.639	0.765	0.320	0.781
0.6	0.95	0.85	0.444	0.360	0.285	0.517	0.808	0.288	0.816
0.6	0.90	0.90	0.603	0.537	0.467	0.977	0.810	0.395	0.826
0.6	0.95	0.90	0.549	0.475	0.401	0.790	0.855	0.356	0.864
0.7	0.90	0.80	0.541	0.445	0.349	0.749	0.720	0.260	0.735
0.7	0.95	0.80	0.494	0.390	0.294	0.606	0.760	0.233	0.768
0.7	0.90	0.85	0.635	0.559	0.474	1.136	0.765	0.320	0.781
0.7	0.95	0.85	0.587	0.500	0.408	0.918	0.808	0.288	0.816
0.7	0.90	0.90	0.726	0.673	0.612	1.737	0.810	0.395	0.826
0.7	0.95	0.90	0.681	0.616	0.543	1.404	0.855	0.356	0.864

f_u is from (6), f'_1 is from (10) and Ψ is probability (11) when $\delta_{\text{J}} = \delta_{\text{NJ}} = \delta$.

DOI: 10.1002/pst

(8) to hold. That is, when both π and $1-\beta'$ are large, showing $\hat{\delta}_J > \pi\delta$ may actually require a larger sample size for Japanese patients than that for showing the overall significant treatment effect. In addition, combining data from a study with Japanese patients' participation and a study without Japanese patients' participation may not increase the chance of showing consistency in treatment effect particularly if the effect of the latter study is larger.

To have a positive trial and simultaneously satisfy the MHLW requirement, instead of considering (3) in isolation, we may consider the probability

$$\Pr(\hat{\delta}_J - \pi\hat{\delta}_{all} > 0, \hat{\delta}_{all} - z_{1-\alpha/2}\sigma/\sqrt{N/2} > 0 | \delta_J, \delta_{NJ}) \tag{11}$$

The covariance between $\hat{\theta} = N(\hat{\delta}_J - \pi\hat{\delta}_{all}) = (N - \pi N_J)\hat{\delta}_J - \pi N_{NJ}\hat{\delta}_{NJ}$ and $N\hat{\delta}_{all}$ is $2(1-\pi)\sigma^2 N$ and the correlation can be calculated as

$$\rho = \frac{(1-\pi)\sqrt{N_J}}{\sqrt{N + (\pi^2 - 2\pi)N_J}}$$

If N_J is replaced by $f_1 N$ in the above formula,

$$\rho = \frac{z_{1-\beta'}}{z_{1-\alpha/2} + z_{1-\beta}}$$

depends only on α, β and β', and is independent of π. This is also clearly demonstrated in Table I. When $\delta_J = \delta_{NJ} = \delta$, (11) is

$$\Psi = \Pr\left(Z_1 > -\rho\sqrt{\frac{N}{2}}\frac{\delta}{\sigma}, Z_2 > z_{1-\alpha/2} - \sqrt{\frac{N}{2}}\frac{\delta}{\sigma}\right)$$

where (Z_1, Z_2) has a bivariate standard normal distribution with correlation ρ. For any fixed a and b,

$$\Pr(Z_1 > a, Z_2 > b) = \int_b^\infty \varphi(Z_2)\Phi\left(\frac{\rho Z_2 - a}{\sqrt{1-\rho^2}}\right) dZ_2$$

is an increasing function of ρ, where φ and Φ are the density and cumulative distribution of the standard normal distribution, respectively. The minimum is achieved when $\rho = 0$. There is no closed-form solution for the minimum number of Japanese patients such that (11) is greater than a pre-specified value even given a fixed N. Numerical computation has to be used for finding the minimum number. For N of overall power $1-\beta$ and N_{1J} of $1-\beta'$, (11) is at least $(1-\beta)(1-\beta')$. Table I provides more precise values of (11) for various parameter configurations. Since the correlations are all not very strong as given in Table I, values of Ψ are only slightly higher than those of $(1-\beta)(1-\beta')$. Obviously, if the overall sample size is for $1-\beta$ power and the fraction of the Japanese sample size ensures the probability in (3) to be no less than $\Psi/(1-\beta)$, (11) will be very close to Ψ. On the other hand, since

$$\begin{aligned}
&\Pr(\hat{\delta}_J > \pi\hat{\delta}_{all} | \text{significant overall effect}) \\
&= \Pr(\hat{\delta}_J - \pi\hat{\delta}_{all} > 0 | \hat{\delta}_{all} - z_{1-\alpha/2}\sigma/\sqrt{N/2} > 0) \\
&= \frac{\Pr(\hat{\delta}_J - \pi\hat{\delta}_{all} > 0, \hat{\delta}_{all} - z_{1-\alpha/2}\sigma/\sqrt{N/2} > 0)}{\Pr(\hat{\delta}_{all} - z_{1-\alpha/2}\sigma/\sqrt{N/2} > 0)}
\end{aligned} \tag{12}$$

the Japanese sample size fraction f_1 also ensures the conditional probability (12) to be close to $1-\beta'$ (note that the numerator is Ψ and the denominator is the overall power $1-\beta$ when $\delta_J = \delta_{NJ} = \delta$). For example, for $\pi = 0.5$, $1-\beta = 0.95$ and N_{1J} of $1-\beta' = 0.8$, Ψ is 0.768 and (12) is $0.768/0.95 = 0.808$, very close to $1-\beta' = 0.8$. The advantage of using f_1 to approximate the required Japanese sample size fraction for (12) to be $1-\beta'$ is its simplicity. There is a closed-form solution for f_1.

We sometimes use an unbalanced design for a clinical trial. That is, the sample sizes for the placebo group and the treatment group are different. Suppose that the sample size for the placebo group is N and the sample size for the treatment group is kN. Then we just need to replace $2\sigma^2$ with $((k+1)/k)\sigma^2$ in all formulas presented above for the calculations. After obtaining the sample size for Japanese patients N_J for the placebo group, the sample size for Japanese patients for the treatment group is simply kN_J. Since the formulas for f_u and f_1' are independent of $((k+1)/k)\sigma^2$, f_u and f_1' remain the same no matter whether we use a balanced or an unbalanced design. The formulas above can also be extended to the case of a binary endpoint. We just need to change $2\sigma^2$ to $p_1(1-p_1) + p_0(1-p_0)$, where p_1

DOI: 10.1002/pst

and p_0 are the event rates for the binary endpoint for treatment and placebo, respectively.

3. SURVIVAL ENDPOINT

For a survival endpoint, the following proportional hazards model is often considered:

$H{:}\lambda_1(t) = \lambda_0(t)e^{\gamma}$

where $\lambda_1(t)$ is the hazard function for treatment, $\lambda_0(t)$ is the hazard function for placebo and e^{γ} is the hazard ratio between treatment and placebo. The power calculation is often based on the log-rank test

$$T \sim N(\mu, 1) \tag{13}$$

where

$$\mu = \gamma\sqrt{E}/2 \tag{14}$$

and E is the expected total number of events from the two groups combined. For a given significance level α (two-sided test) and power $1-\beta$, the required expected total number of events from the two groups is

$$E = \frac{4(z_{1-\alpha/2} + z_{1-\beta})^2}{\gamma^2} \tag{15}$$

From (13) and (14),

$$\hat{\gamma} = \frac{2T}{\sqrt{\hat{E}}} \sim N\left(\gamma, \frac{4}{E}\right) \tag{16}$$

where $\hat{E}$ is the observed value of E. Therefore, $\hat{\gamma}$ is an estimate of γ, $e^{\hat{\gamma}}$ is the estimate of the relative risk and $(1-e^{\hat{\gamma}})$ is the estimate of the risk reduction for the treatment. When there is a positive treatment effect, $\hat{\gamma}$ should be less than 0 and $e^{\hat{\gamma}}$ should be less than 1 (the smaller the better). There are two approaches for applying (1) to a survival endpoint. Since γ is supposed to be negative, the first approach is to directly consider the probability of $\hat{\gamma}_J/\hat{\gamma}_{all} > \pi$ or $\hat{\gamma}_J < \pi\hat{\gamma}_{all}$, where $\hat{\gamma}_J$ is the estimate of γ based on the data from Japanese patients and $\hat{\gamma}_{all}$ is the corresponding estimate for the entire group. The second approach is to consider the probability of $(1-e^{\hat{\gamma}_J})/(1-e^{\hat{\gamma}_{all}}) > \pi$ or $e^{\hat{\gamma}_J} < (1-\pi) + \pi e^{\hat{\gamma}_{all}}$ and $\hat{\gamma}_J < \log(1-\pi+\pi e^{\hat{\gamma}_{all}})$. Note that $f(x) = \pi x - \log(1-\pi+\pi e^{x})$ is an increasing function for negative x and $f(0) = 0$. Hence, for negative $\hat{\gamma}_{all}$, $\pi\hat{\gamma}_{all} < \log(1-\pi+\pi e^{\hat{\gamma}_{all}})$ and very likely

$$\Pr\left(\frac{1-e^{\hat{\gamma}_J}}{1-e^{\hat{\gamma}_{all}}} > \pi | \gamma_J, \gamma_{all}\right) \geqslant \Pr\left(\frac{\hat{\gamma}_J}{\hat{\gamma}_{all}} > \pi | \gamma_J, \gamma_{all}\right) \tag{17}$$

for negative γ_J and γ_{all}, the true values corresponding to $\hat{\gamma}_J$ and $\hat{\gamma}_{all}$, respectively. Because risk reduction is often used for measuring the treatment effect for a survival endpoint, in the following, (1) will be interpreted as

$$\Pr\left(\frac{1-e^{\hat{\gamma}_J}}{1-e^{\hat{\gamma}_{all}}} > \pi \middle| \gamma_J, \gamma_{all}\right) \geqslant 1-\beta' \tag{18}$$

for a survival endpoint. Since it will be difficult to derive the correlation between $\hat{\gamma}_J$ and $\hat{\gamma}_{all}$ if the pooled data are used to derive $\hat{\gamma}_{all}$, we will consider $\hat{\gamma}_{all}$ as a weighted combination of $\hat{\gamma}_J$ and $\hat{\gamma}_{NJ}$ in the following derivations, where $\hat{\gamma}_{NJ}$ is the estimate of the true γ_{NJ}, the γ for non-Japanese patients. That is,

$$\hat{\gamma}_{all} = w\hat{\gamma}_J + (1-w)\hat{\gamma}_{NJ}, \quad 0 \leqslant w \leqslant 1 \tag{19}$$

Asymptotically,

$$\begin{pmatrix} \hat{\gamma}_J \\ \hat{\gamma}_{all} \end{pmatrix} \sim N\left(\begin{pmatrix} \gamma_J \\ \gamma_{all} \end{pmatrix}, \Sigma\right)$$

where $\gamma_{all} = w\gamma_J + (1-w)\gamma_{NJ}$ and the covariance matrix

$$\Sigma = \mathrm{Var}\left(\begin{pmatrix} \hat{\gamma}_J \\ \hat{\gamma}_{all} \end{pmatrix}\right) = \begin{pmatrix} \frac{4}{E_J} & w\frac{4}{E_J} \\ w\frac{4}{E_J} & w^2\frac{4}{E_J} + (1-w)^2\frac{4}{E_{NJ}} \end{pmatrix}$$

where E_J is the expected number of events from Japanese patients and E_{NJ} is the expected number of events from non-Japanese patients. Two types of weight strategies could be considered. One is to use the inverse of variance or the number of events $w = E_J/E$ as the weight (similar to using sample size as the weight for normal endpoint). This strategy will give the variance of $\hat{\gamma}_{all}$ as $4/E$, the same variance as that from a pooled analysis.

DOI: 10.1002/pst

Another strategy is to use $w = 0$ as the weight. This basically assesses the consistency of treatment effects between Japanese patients and non-Japanese patients. It can be applied to bridging study as well. Evidently, if the treatment effect for Japanese patients is consistent with that of non-Japanese patients, the effect for Japanese patients should also be consistent with the overall effect. Because $\hat{\gamma}_{\mathrm{J}}$ and $\hat{\gamma}_{\mathrm{NJ}}$ are independent, the sample size calculation for Japanese patients based on $w = 0$ could be conceptually simpler. In case some practitioners do want to use such an approach, we provide the calculation based on $w = 0$ here for completeness. As we will see subsequently without surprise, the sample size for Japanese patients obtained in such a way is consistently much larger than the one based on $w = E_{\mathrm{J}}/E$. In the following, two delta methods will be used to derive the required number of events for Japanese patients. Within each delta method, three approaches are considered: first for $w = E_{\mathrm{J}}/E$, second for $w = 0$ and third for treating $\hat{\gamma}_{\mathrm{all}}$ as a fixed constant $\gamma_{\mathrm{all}} = \gamma$.

3.1. Delta method 1

One delta method is based on the asymptotic distribution

$$\hat{\phi} = \pi e^{\hat{\gamma}_{\mathrm{all}}} - e^{\hat{\gamma}_{\mathrm{J}}} \sim N(\phi, \varsigma^2)$$

with $\phi = \pi e^{\gamma_{\mathrm{all}}} - e^{\gamma_{\mathrm{J}}}$ and

$$\varsigma^2 = \frac{4}{E_{\mathrm{J}}}(e^{\gamma_{\mathrm{J}}})^2 - \frac{8w}{E_{\mathrm{J}}}\pi(e^{\gamma_{\mathrm{J}}})(e^{\gamma_{\mathrm{all}}}) + \left(w^2\frac{4}{E_{\mathrm{J}}} + (1-w)^2\frac{4}{E_{\mathrm{NJ}}}\right)\pi^2(e^{\gamma_{\mathrm{all}}})^2$$

Then, (18) becomes

$$\frac{1-\pi+\phi}{\varsigma} > z_{1-\beta'} \tag{20}$$

As for the case of a normal endpoint, suppose that $\gamma_{\mathrm{J}} = u\gamma_{\mathrm{NJ}}$, where u again is around 1. Let g_u be the corresponding minimum fraction of the number of events for Japanese patients satisfying (20). Then, $E_{u\mathrm{J}}=g_uE$ and $E_{u\mathrm{NJ}}=(1-g_u)E$. If $w = g_u$,

$$\varsigma^2 = \frac{4}{E}\left[\frac{1}{g_u}e^{2u\gamma_{\mathrm{NJ}}} - 2\pi e^{((1+g_u)u+1-g_u)\gamma_{\mathrm{NJ}}} + \pi^2 e^{(2ug_u+2(1-g_u))\gamma_{\mathrm{NJ}}}\right]$$

and g_u should satisfy

$$\frac{\sqrt{E}(1-\pi+\pi e^{(ug_u+1-g_u)\gamma_{\mathrm{NJ}}}-e^{u\gamma_{\mathrm{NJ}}})}{2\sqrt{(1/g_u)e^{2u\gamma_{\mathrm{NJ}}}-2\pi e^{(u(1+g_u)+1-g_u)\gamma_{\mathrm{NJ}}}+\pi^2 e^{2(ug_u+1-g_u)\gamma_{\mathrm{NJ}}}}} = z_{1-\beta'}$$

where E is given by (15) or

$$E = \frac{4(z_{1-\alpha/2}+z_{1-\beta})^2}{(ug_u+1-g_u)^2\gamma_{\mathrm{NJ}}^2}$$

For a general u, there is no closed-form solution for g_u. For given α, β, β', π, u and γ_{NJ}, a numerical approach should be used to derive the solution. However, when $u = 1$, or $\gamma_{\mathrm{J}} = \gamma_{\mathrm{NJ}} = \gamma$, a closed-form solution for g_u exists:

$$g_1 = \frac{4e^{2\gamma}z_{1-\beta'}^2}{E(1-\pi)^2(1-e^{\gamma})^2 + 4e^{2\gamma}(2\pi-\pi^2)z_{1-\beta'}^2}$$

The number of events for Japanese patients in this case is denoted as

$$E_{1\mathrm{J}} = g_1E \tag{21}$$

For $\gamma_{\mathrm{J}} = \gamma_{\mathrm{all}} = \gamma$ and weight $w = 0$, (20) becomes

$$\frac{(1-\pi)(1-e^{\gamma})}{2e^{\gamma}\sqrt{1/E_{\mathrm{J}}+\pi^2/(E-E_{\mathrm{J}})}} > z_{1-\beta'}$$

or

$$E_{\mathrm{J}}^2 - \left(E + \frac{4e^{2\gamma}z_{1-\beta'}^2(1+\pi)}{(1-\pi)(1-e^{\gamma})^2}\right)E_{\mathrm{J}} + \frac{4e^{2\gamma}z_{1-\beta'}^2}{(1-\pi)^2(1-e^{\gamma})^2}E < 0 \tag{22}$$

DOI: 10.1002/pst

The lower bound of the solution for (22) is

$$E_{\mathrm{J}} > \frac{1}{2}\left(E+\frac{4e^{2\gamma}z_{1-\beta'}^2(1+\pi)}{(1-\pi)(1-e^{\gamma})^2} - \sqrt{\left(E+\frac{4e^{2\gamma}z_{1-\beta'}^2(1+\pi)}{(1-\pi)(1-e^{\gamma})^2}\right)^2 - \frac{16e^{2\gamma}z_{1-\beta'}^2 E}{(1-\pi)^2(1-e^{\gamma})^2}}\right) = E_{2\mathrm{J}} \quad (23)$$

As for the case of a normal endpoint, for convenience, some practitioners may replace $\hat{\gamma}_{\mathrm{all}}$ by γ_{all} in (18) in the calculation. Under $\gamma_{\mathrm{J}} = \gamma_{\mathrm{all}} = \gamma$, the solution for

$$\Pr\left(\frac{1-e^{\hat{\gamma}_{\mathrm{J}}}}{1-e^{\gamma_{\mathrm{all}}}} > \pi \middle| \gamma_{\mathrm{J}} = \gamma_{\mathrm{all}} = \gamma\right) \geqslant 1-\beta'$$

or

$$\Pr(\hat{\gamma}_{\mathrm{J}} < \log(1-\pi(1-e^{\gamma}))|\gamma_{\mathrm{J}} = \gamma_{\mathrm{all}} = \gamma) \geqslant 1-\beta'$$

is

$$E_J > \frac{4z_{1-\beta'}^2}{(\gamma - \log(1-\pi(1-e^{\gamma})))^2} = E_{3\mathrm{J}} \quad (24)$$

3.2. Delta Method 2

Similar to Hung *et al.* [11] for the case of a non-inferiority assessment, another delta method is based on the asymptotic distribution,

$$\hat{\eta} = \log\frac{1-e^{\hat{\gamma}_{\mathrm{J}}}}{1-e^{\hat{\gamma}_{\mathrm{all}}}} \sim N(\eta, \upsilon^2)$$

where

$$\eta = \log\frac{1-e^{\gamma_{\mathrm{J}}}}{1-e^{\gamma_{\mathrm{all}}}}$$

and

$$\upsilon^2 = \frac{4}{E_{\mathrm{J}}}\left(\frac{e^{\gamma_{\mathrm{J}}}}{1-e^{\gamma_{\mathrm{J}}}}\right)^2 - \frac{8w}{E_{\mathrm{J}}}\left(\frac{e^{\gamma_{\mathrm{J}}}}{1-e^{\gamma_{\mathrm{J}}}}\right)\left(\frac{e^{\gamma_{\mathrm{all}}}}{1-e^{\gamma_{\mathrm{all}}}}\right) + \left(w^2\frac{4}{E_{\mathrm{J}}} + (1-w)^2\frac{4}{E_{\mathrm{NJ}}}\right)\left(\frac{e^{\gamma_{\mathrm{all}}}}{1-e^{\gamma_{\mathrm{all}}}}\right)^2$$

When $\gamma_{\mathrm{J}} = \gamma_{\mathrm{NJ}} = \gamma$, $\eta = 0$,

$$\upsilon^2 = 4\left(\frac{e^{\gamma}}{1-e^{\gamma}}\right)^2\left(\frac{1}{E_{\mathrm{J}}} - \frac{2w}{E_{\mathrm{J}}} + w^2\frac{1}{E_{\mathrm{J}}} + (1-w)^2\frac{1}{E_{\mathrm{NJ}}}\right) = 4\left(\frac{e^{\gamma}}{1-e^{\gamma}}\right)^2(1-w)^2\left(\frac{1}{E_{\mathrm{J}}} + \frac{1}{E_{\mathrm{NJ}}}\right)$$

(18) becomes

$$\Pr(\hat{\eta} > \log\pi|\gamma_{\mathrm{J}} = \gamma_{\mathrm{NJ}} = \gamma) = \Pr\left[Z > \frac{\log\pi}{\upsilon}\right] \geqslant 1-\beta'$$

That is,

$$\frac{-\log\pi}{\upsilon} \geqslant z_{1-\beta'} \quad (25)$$

When $w = E_{\mathrm{J}}/E$, (25) becomes

$$E_{\mathrm{J}} > \frac{4e^{2\gamma}z_{1-\beta'}^2 E}{E(\log\pi)^2(1-e^{\gamma})^2 + z_{1-\beta'}^2 4e^{2\gamma}} = E_{4\mathrm{J}} \quad (26)$$

Since $2\pi - \pi^2 \leqslant 1$ and $1-\pi \leqslant -\log\pi$ if $0 < \pi \leqslant 1$, $E_{4\mathrm{J}} \leqslant E_{1\mathrm{J}}$ always holds as long as $0 < \pi \leqslant 1$. When $w = 0$, (25) becomes

$$E_{5\mathrm{J}} = \frac{1}{2}\left(E - \sqrt{E^2 - \frac{16e^{2\gamma}z_{1-\beta'}^2 E}{(\log\pi)^2(1-e^{\gamma})^2}}\right) < E_{\mathrm{J}} < \frac{1}{2}\left(E + \sqrt{E^2 - \frac{16e^{2\gamma}z_{1-\beta'}^2 E}{(\log\pi)^2(1-e^{\gamma})^2}}\right) \quad (27)$$

We just need the lower bound. The upper bound is for preventing E_{J} becoming too close to E and ensuring a sufficient number E_{NJ} of events from non-Japanese patients for (18) to hold when $w = 0$. If we treat $\hat{\gamma}_{\mathrm{all}}$ as γ_{all} in $\hat{\eta}$, under $\gamma_{\mathrm{J}} = \gamma_{\mathrm{all}} = \gamma$, we have

$$E_{\mathrm{J}} > \frac{4e^{2\gamma}z_{1-\beta'}^2}{(\log\pi)^2(1-e^{\gamma})^2} = E_{6\mathrm{J}} > E_{4\mathrm{J}} \quad (28)$$

Table II provides results for comparing $E_{i\mathrm{J}}$, $i = 1, 2, \ldots, 6$ (see (21), (23), (24) and (26)–(28)), with the overall number of events E for different π, $1-e^{\gamma}$, $1-\beta$ and $1-\beta'$. There are big differences between the numbers of events for Japanese patients between the two delta methods. For example, when $\alpha = 0.05$, $\pi = 0.5$, $1-e^{\gamma} = 0.2$, $1-\beta = 0.9$ and $1-\beta' = 0.8$, $E_{1\mathrm{J}} = 156$ from the first delta method is 18.5% of $E = 844$. However, $E_{4\mathrm{J}}$ from the second delta method is only 10.1% of E.

DOI: 10.1002/pst

Table II. Number of events for survival endpoint ($\alpha = 0.05$).

π	$1-e^{\gamma}$	$1-\beta$	$1-\beta'$	E	E_{1J}	E_{2J}	E_{3J}	E_{4J}	E_{5J}	E_{6J}
0.5	0.2	0.90	0.80	844	156	195	204	85	108	94
0.5	0.2	0.95	0.80	1044	160	192	204	87	105	94
0.5	0.3	0.90	0.80	330	54	66	75	29	36	32
0.5	0.3	0.95	0.80	409	55	65	75	30	35	32
0.5	0.4	0.90	0.80	161	23	27	34	12	15	13
0.5	0.4	0.95	0.80	199	23	26	34	12	14	13
0.5	0.2	0.90	0.85	844	221	316	310	122	183	143
0.5	0.2	0.95	0.85	1044	230	303	310	126	171	143
0.5	0.3	0.90	0.85	330	77	104	114	42	59	49
0.5	0.3	0.95	0.85	409	80	101	114	44	57	49
0.5	0.4	0.90	0.85	161	33	42	52	18	24	20
0.5	0.4	0.95	0.85	199	34	41	52	18	23	20
0.6	0.2	0.90	0.80	844	221	359	312	144	245	174
0.6	0.2	0.95	0.80	1044	231	331	312	149	220	174
0.6	0.3	0.90	0.80	330	77	115	113	50	77	59
0.6	0.3	0.95	0.80	409	80	109	113	52	72	59
0.6	0.4	0.90	0.80	161	33	45	51	21	30	24
0.6	0.4	0.95	0.80	199	34	44	51	22	29	24
0.6	0.2	0.90	0.85	844	301	—	473	201	—	263
0.6	0.2	0.95	0.85	1044	319	—	473	210	—	263
0.6	0.3	0.90	0.85	330	107	—	172	71	—	90
0.6	0.3	0.95	0.85	409	112	194	172	74	133	90
0.6	0.4	0.90	0.85	161	46	84	77	30	58	37
0.6	0.4	0.95	0.85	199	48	73	77	31	49	37

See (21), (23), (24) and (26)–(28) for the definitions of E_{iJ}, $i = 1, 2, \ldots, 6$.

Simulations were conducted to compare the performances of the two delta methods and the results will be presented in Section 5. For the scenarios of large π, small $1-e^{\gamma}$ and high $1-\beta'$, the numbers of events for Japanese patients obtained based on $w = 0$ for (18) to hold do not exist no matter whether the first or the second delta method is used (see the missing values for E_{2J} and E_{5J} in Table II).

4. TRANSFORMING THE NUMBER OF EVENTS INTO SAMPLE SIZE FOR A SURVIVAL ENDPOINT

Before we start the simulation, we first transform the number of events into sample size assuming that event time and censoring time follow exponential distributions. As in Quan *et al.* [12], two types of trial designs are considered. One is the fixed stopping time design, in which even though patients enter into the study at staggered calendar time, they all stop treatment at the same calendar time once certain total number of events is reached. This design is often used for long-term overall survival trials. Another is the fixed study duration design, in which all patients have the same fixed study duration no matter when they enter into the study.

For the fixed stopping time design, given the event hazard rate λ_0 for placebo and λ_1 for treatment, the common discontinuation hazard rate of τ ($\tau = 0$ for intent-to-treat overall survival endpoint), a constant accrual rate of r patients per time unit per group for an A-time-unit accrual period, F-time-unit follow-up after the last randomization and $L = F + A$, a patient entering the study at time s will have a probability of $\Pr(X_i \leqslant \min(L - s, C))$ to have an event during the study,

DOI: 10.1002/pst

where X_i is the event time for the treatment i and C is the discontinuation time and is independent with X_i. Since

$$\begin{aligned}
&\Pr(X_i \leqslant \min(L-s, C)) \\
&= \int_0^\infty \Pr(X_i \leqslant L-s, X_i \leqslant C | C=c)\tau e^{-\tau c}\,\mathrm{d}c \\
&= \int_0^{L-s} \Pr(X_i \leqslant C | C=c)\tau e^{-\tau c}\,\mathrm{d}c \\
&\quad + \int_{L-s}^\infty \Pr(X_i \leqslant L-s | C=c)\tau e^{-\tau c}\,\mathrm{d}c \\
&= \int_0^{L-s} (1-e^{-\lambda_i c})\tau e^{-\tau c}\,\mathrm{d}c \\
&\quad + \int_{L-s}^\infty (1-e^{-\lambda_i (L-s)})\tau e^{-\tau c}\,\mathrm{d}c \\
&= \frac{\lambda_i}{\lambda_i+\tau}(1-e^{-(\lambda_i+\tau)(L-s)}) = H_i(s)
\end{aligned}$$

the expected number of events for the treatment i is

$$\begin{aligned}
&\int_0^A r \Pr(X_i \leqslant \min(L-s, C))\,\mathrm{d}s \\
&= \int_0^A r H_i(s)\,\mathrm{d}s \\
&= \frac{r\lambda_i}{\lambda_i+\tau}\left(A - \frac{e^{-(\lambda_i+\tau)L}}{\lambda_i+\tau}(e^{(\lambda_i+\tau)A}-1)\right) \\
&= rV_i
\end{aligned}$$

Thus, the total number of events combining from the two groups is

$$E = E_0 + E_1 = r(V_0 + V_1)$$

Then, the total sample size per group can be derived as

$$N = rA = AE/(V_0 + V_1)$$

For the fixed study duration design, suppose that the length of the study duration for all patients is L. Using the above notations, the total number of events combining from the two groups is

$$E = E_0 + E_1 = N(H_0(0) + H_1(0))$$

The total sample size per group is

$$N = E/(H_0(0) + H_1(0))$$

The number of Japanese patients can be derived by replacing E with E_J in the two formulas. When we use the number of patients instead of number of events for Japanese patients, for the fixed stopping time design, we have to ensure that the Japanese sites are opened earlier or simultaneously with sites in other regions. If it is anticipated that Japanese sites are opened later than other sites, we have to increase the Japanese sample size to ensure a sufficient number of events for Japanese patients when the trial reaches the total number of E events.

5. SIMULATION RESULTS

Simulation is used to compare the performance of the two delta methods. For the simulation, we only consider the fixed study duration design ($A = 0$). Results for the fixed stopping time design should be similar. Besides those parameters used for Table II, we set $\lambda_0 = 5\%$, $\lambda_1 = \lambda_0 e^{\gamma}$, $\tau = 0$ (no discontinuation for intent-to-treat overall survival analysis) and $L = 36$. Based on these assumptions, the overall sample size per group can be derived using $N = E/(H_0(0) + H_1(0))$ and the sample size for Japanese patients per treatment group can be derived similarly using $N^*_{iJ} = E_{iJ}/(H_0(0) + H_1(0))$. With sample size N, event times for individual patients are simulated using the exponential distribution of the corresponding hazard rate and are censored if they are greater than L. Among N survival times, N^*_{iJ} are for Japanese patients. The Cox proportional hazards model is then used to derive the estimate of hazard ratio $\hat{\gamma}_{\text{all}}$ for the entire group and $\hat{\gamma}_{iJ}$ for Japanese patients. The proportion P_{iJ} of $(1-e^{\hat{\gamma}_{iJ}})/(1-e^{\hat{\gamma}_{\text{all}}}) > \pi$ corresponding to N^*_{iJ} based on 10 000 runs is the estimate of probability $\Pr((1-e^{\hat{\gamma}_{iJ}})/(1-e^{\hat{\gamma}_{\text{all}}}) > \pi | \gamma_{\text{all}} = \gamma_{iJ} = \gamma)$, which should be compared with $1-\beta'$ to assess the performance of the two delta methods and the three approaches within each delta method. Results are presented in Table III. Values of P_{1J}, the estimates of the probability based on E_{1J}, are slightly smaller than the nominal value. Values of P_{3J} always exceed the nominal values, while those of P_{2J} exceed the nominal values most of the times. Even though the number of events in Japanese patients (E_{4J}, E_{5J} and E_{6J}) and the corresponding samples size based on the second delta method are substantially smaller than those of the first delta method, values of P_{4J}, P_{5J} and P_{6J} are not much

DOI: 10.1002/pst

Table III. Simulation results for survival endpoint based on 10 000 runs ($\alpha = 0.05$, $\lambda_0 = 5\%$, $\tau = 0$ and $L = 36$).

π	$1-e^{\gamma}$	$1-\beta$	$1-\beta'$	E	P_{1J}	P_{2J}	P_{3J}	P_{4J}	P_{5J}	P_{6J}
0.5	0.2	0.90	0.80	844	0.779	0.817	0.826	0.710	0.741	0.727
0.5	0.2	0.95	0.80	1044	0.780	0.803	0.814	0.715	0.735	0.726
0.5	0.3	0.90	0.80	330	0.771	0.792	0.815	0.702	0.730	0.717
0.5	0.3	0.95	0.80	409	0.774	0.794	0.808	0.706	0.729	0.714
0.5	0.4	0.90	0.80	161	0.762	0.779	0.816	0.698	0.722	0.711
0.5	0.4	0.95	0.80	199	0.775	0.781	0.808	0.698	0.724	0.713
0.5	0.2	0.90	0.85	844	0.836	0.892	0.890	0.757	0.811	0.779
0.5	0.2	0.95	0.85	1044	0.837	0.881	0.883	0.755	0.799	0.773
0.5	0.3	0.90	0.85	330	0.819	0.865	0.880	0.745	0.784	0.759
0.5	0.3	0.95	0.85	409	0.819	0.854	0.873	0.748	0.778	0.765
0.5	0.4	0.90	0.85	161	0.815	0.849	0.877	0.735	0.769	0.739
0.5	0.4	0.95	0.85	199	0.818	0.846	0.872	0.739	0.766	0.750
0.6	0.2	0.90	0.80	844	0.791	0.872	0.842	0.726	0.807	0.752
0.6	0.2	0.95	0.80	1044	0.789	0.839	0.831	0.732	0.781	0.751
0.6	0.3	0.90	0.80	330	0.782	0.846	0.842	0.731	0.782	0.752
0.6	0.3	0.95	0.80	409	0.781	0.826	0.832	0.725	0.770	0.741
0.6	0.4	0.90	0.80	161	0.773	0.817	0.839	0.717	0.759	0.733
0.6	0.4	0.95	0.80	199	0.775	0.806	0.827	0.725	0.747	0.732
0.6	0.2	0.90	0.85	844	0.835	—	0.919	0.777	—	0.816
0.6	0.2	0.95	0.85	1044	0.839	—	0.906	0.777	—	0.812
0.6	0.3	0.90	0.85	330	0.833	—	0.911	0.771	—	0.808
0.6	0.3	0.95	0.85	409	0.828	0.917	0.898	0.772	0.855	0.798
0.6	0.4	0.90	0.85	161	0.829	0.918	0.904	0.765	0.862	0.793
0.6	0.4	0.95	0.85	199	0.818	0.883	0.890	0.765	0.824	0.788

P_{iJ} is the probability in (18) for E_{iJ} events from Japanese patients when $\gamma_J = \gamma_{NJ}$.

smaller than the nominal level of $1-\beta'$. Combining these findings with the results from Table II, E_{1J} is a reasonable choice for determining the number of events for Japanese patients.

6. EXAMPLES

Two examples are used to illustrate the application of the approaches discussed in this paper, one for a normal endpoint and the other for a survival endpoint.

6.1. Example of a normal endpoint

A multi-regional trial was designed to evaluate the effect of an investigational drug on change from baseline in HbA_{1c}. In order to have more safety data for the investigational drug, the trial used an unbalanced design with a 2:1 ratio for the active treatment and placebo groups and the trial was over-powered for the primary efficacy endpoint. Basically, to detect a difference of 0.5% in change from baseline in HbA_{1c} between the two groups, 372 patients for the active treatment group and 186 patients for the placebo group for the entire study provided an overall power of 99%, assuming the common standard deviation of 1.3% with a two-sided test at the 5% significance level. With these specifications, we first calculate f_1 and $f_1^{'}$ using formulas (7) and (10) and the results are provided in Table IV. As mentioned earlier, they remain the same for both balanced and unbalanced designs. Because of the high power, they are smaller than those presented in Table I. We then multiply the overall sample size by f_1 or $f_1^{'}$ to get the sample size for Japanese patients. For $\pi = 0.5$ and $1-\beta' = 0.8$, the sample sizes for Japanese patients are 26 and 51 (or 52 to satisfy the 2:1 ratio) for the placebo and active treatment groups, respectively. If $1-\beta'$ is increased from 0.8 to 0.9, the sample sizes need to

DOI: 10.1002/pst

Table IV. Sample size for Japanese patients in an HbA_{1c} trial.

				N_{1J} based on f_1		N'_{1J} based on f'_1	
π	$1-\beta'$	f_1	f'_1	Placebo	Treatment	Placebo	Treatment
0.5	0.80	0.138	0.154	26	51	29	57
0.5	0.85	0.199	0.234	37	74	43	87
0.5	0.90	0.282	0.358	52	105	67	133
0.6	0.80	0.200	0.241	37	75	45	90
0.6	0.85	0.280	0.365	52	104	68	136
0.6	0.90	0.380	0.559	71	141	104	208
0.7	0.80	0.308	0.428	57	115	80	159
0.7	0.85	0.408	0.650	76	152	121	242
0.7	0.90	0.522	0.993	97	194	185	369

f_1 is from (7) and f'_1 is from (10). N_{1J} and N'_{1J} are the corresponding sample sizes for Japanese patients.

be approximately doubled. Sample sizes obtained through f'_1 are much bigger than those through f_1 especially when π is large.

6.2. Example of a survival endpoint

A multi-regional oncology trial was designed to assess the effect of an investigational drug on the primary efficacy endpoint of overall survival. The expected median survival time for the patient population in the control arm was 21 months or $\lambda_0 = 3.30\%$/month. A hazard rate reduction of 20% or $\lambda_1 = 2.64\%$/month in the test arm compared with the control arm was expected. Assuming that survival times were exponentially distributed in both treatment arms, a total of 844 deaths would be needed to detect a hazard rate reduction of 20% with 90% power using a two-sided log-rank test at a significance level of 5%. Based on an anticipated accrual period of 42 months followed by a 12-month follow-up after the randomization of the last patient, $A = 42$, $F = 12$ and $L = 54$ months. Since the overall survival was the primary endpoint, $\tau = 0$ or there was no discontinuation for the intent-to-treat analysis. Under these assumptions, we derive V_0 as 26.71 and V_1 as 23.51. Given $E = 844$, the total sample size for the entire study is 1412. For $\pi = 0.5$, $1-\beta' = 0.8$ and E_{1J}, as given in Table V, the total number of Japanese patients combined from placebo and active treatment is 261. If E_{4J} is used instead, the total number of Japanese patients combined from placebo and active treatment is only 142. However, based on simulation results in Section 5, it provides only approximately 71% probability in (18). If larger values of π and $1-\beta'$ are considered, the required sample size for Japanese patients needs a substantial increase.

7. DISCUSSION

The current trend of global new drug development strategy is moving away from a delayed isolated bridging study approach to a multi-regional trial approach. The advantage of this multi-regional trial strategy is to streamline and expedite the new drug development in all regions. If trial results really demonstrate the consistency of treatment effects across regions and ethnical groups, patients can access effective and safe drugs simultaneously worldwide. Method 1 in the MHLW guidance focuses on ensuring that the treatment effect on Japanese patients is not much less than the overall effect.

When the true treatment effects for Japanese patients and the other patients are slightly different, numerical approaches have to be used to derive the sample size/number of events for Japanese patients. If they are the same, closed-form solutions are available. For a normal endpoint and a power of 90% for the overall sample size, the number of Japanese patients is 22.4% of the overall number of patients to meet the minimum requirement of Method 1 in the guidance. For a survival endpoint, if $\hat{\gamma}_J/\hat{\gamma}_{all} > \pi$ is used to assess the consistency of

DOI: 10.1002/pst

Table V. Number of events for Japanese patients in an oncology trial.

π	$1-\beta'$	E_{1J}	E_{2J}	E_{3J}	E_{4J}	E_{5J}	E_{6J}	N^*_{1J}	N^*_{2J}	N^*_{3J}	N^*_{4J}	N^*_{5J}	N^*_{6J}
0.5	0.80	156	195	204	85	108	94	261	326	342	142	181	158
0.5	0.85	221	316	310	122	183	143	370	529	518	205	305	239
0.5	0.90	306	—	474	174	—	219	512	—	792	291	—	366
0.6	0.80	221	359	312	144	245	174	370	600	522	241	409	291
0.6	0.85	301	—	473	201	—	263	503	—	791	336	—	441
0.6	0.90	397	—	723	273	—	403	664	—	1210	456	—	674
0.7	0.80	326	—	542	251	—	356	546	—	906	419	—	596
0.7	0.85	419	—	822	329	—	540	701	—	1374	551	—	904
0.7	0.90	517	—	1256	418	—	826	865	—	2101	698	—	1382

N^*_{iJ} is the Japanese sample size for E_{iJ} events from Japanese patients.

treatment effects between Japanese patients and the entire group, the minimum percentage of Japanese patients required should also be 22.4%. This can be seen when the weight $w = E_J/E$ is used in the $\hat{\gamma}_{all}$ calculation. Because of inequality (17), using $(1-e^{\hat{\gamma}_J})/(1-e^{\hat{\gamma}_{all}}) > \pi$ for the consistency assessment requires smaller percentages of Japanese patients.

The guidance was issued recently. Before it receives broad acceptance, its practicability should be assessed and examined as we conduct more research and gain experiences through actual trials. In this paper, we focused only on the required sample size for effect consistency evaluation based on one method in the guidance. We hope that this initial investigation will contribute to continued progress in this important area.

ACKNOWLEDGEMENTS

We would like to thank the associate editor and the referees for their very thorough review. Their thoughtful comments and suggestions enhanced the manuscript greatly.

REFERENCES

1. ICH International Conference on Harmonization Tripartite Guidance E5. Ethnic factor in the acceptability of foreign data. *The US Federal Register*, Vol. 83, 1998; pp. 31790–31796.
2. Shih WJ. Clinical trials for drug registrations in Asian-Pacific countries: proposal for a new paradigm from a statistical perspective. *Controlled Clinical Trials* 2001; **22**:357–366.
3. Chow SC, Shao J, Hu OYP. Assessing sensitivity and similarity in bridging studies. *Journal of Biopharmaceutical Statistics* 2002; **12**:385–400.
4. Hsiao CF, Xu JZ, Liu JP. A group sequential approach to evaluation of bridging studies. *Journal of Biopharmaceutical Statistics* 2003; **13**:793–801.
5. Uyama Y, Shibata T, Nagai N, Hanaoka H, Toyoshima S, Mori K. Successful bridging strategy based on ICH E5 guideline for drugs approval in Japan. *Clinical Pharmacology & Therapeutics* 2005; **78**:102–113.
6. Ministry of Health, Labour and Welfare of Japan. *Basic Principles on Global Clinical Trials*, September 28, 2007.
7. Kawai N, Chuang-Stein C, Komiyama O, Li Y. An approach to rationalize partitioning sample size into individual regions in a multiregional trial. *Drug Information Journal* 2007; **42**:139–147.
8. Uesaka H. Statistical consideration for multi-regional clinical trials in simultaneous development among 3 regions. *Proceedings of the Biostatistic Forum of Kurume University*, Fukuoka, Japan, September 30, 2006 (in Japanese).
9. Uesaka H. Consistency of results among regions and sample size in a multi-regional clinical trial. *Proceedings of the Symposium on use of Multiple Decision Procedures in R&D*, Kobe, Japan, February 27, 2007 (in Japanese).
10. Sekiguchi R, Ogawa S, Uesaka H. Sample size determination of Japanese patients for multi-regional clinical trial (MRCT) in oncology. Presentation at *Joint Statistical Meetings*, Salt Lake City, USA, July 29–August 2, 2007.
11. Hung HMJ, Wang SJ, Tsong Y, Lawrence J, O'Neil RT. Some fundamental issues with non-inferiority testing in active controlled trials. *Statistics in Medicine* 2003; **22**:213–225.
12. Quan H, Sun Q, Zhang J, Shih WJ. Comparisons between ITT and treatment emergent adverse event analyses. *Statistics in Medicine* 2008; **27**:5356–5376.

DOI: 10.1002/pst