Statistics and Its Interface Volume 17 (2024) 63–68

Aligning sample size calculations with estimands in clinical trials with time-to-event outcomes

Yixin Fang*, Man Jin, and Chengqing Wu

The ICH E9(R1) guidance recommended a framework to align planning, design, conduct, analysis, and interpretation of any clincial trial with its objective and estimand. How to handle intercurrent events (ICEs) is one of the five attributes of an estimand and sample size calculation is a key step in the trial planning and design. Therefore, sample size calculation should be aligned with the estimand and, in particular, with how the ICEs are handled. ICH E9(R1) summarized five strategies for handling ICEs, and five approaches have been proposed in the literature for sample size calculation when planning trials with quantitative and binary outcomes. In this paper, we discuss how to apply the five strategies to deal with ICEs in clinical trials with time-to-event outcomes and propose five approaches for sample size calculation that are aligned with the five strategies, respectively.

AMS 2000 subject classifications: Primary 62N03, 62G10; secondary 62G99.
Keywords and phrases: Clinical trials, Estimand, Intercurrent events, Sample size, Time-to-event.

1. INTRODUCTION

The International Council for Harmonisation (ICH) E9(R1) guidance [7] stated that "central questions for drug development and licensing are to establish the existence, and to estimate the magnitude, of treatment effects: how the outcome of treatment compares to what would have happened to the same subjects under alternative treatment." ICH E9(R1) emphasized that "precision in describing a treatment effect of interest is facilitated by constructing the estimand" and recommended "a framework to align planning, design, conduct, analysis, and interpretation" with the trial estimand. Since sample size calculation is a key step in the planning and design stages of any clinical trial, we should align sample size calculation with the estimand.

As summarized in ICH E9(R1), the five attributes of an estimand are treatment, population, outcome variable, a population-level summary, and how to deal with intercurrent events (ICEs). ICH E9(R1) defined ICEs as "events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest". ICH E9(R1) proposed five strategies to deal with ICEs: (i) treatment-policy strategy, (ii) hypothetical strategy, (iii) composite-variable strategy, (iv) while-on-treatment strategy, and (v) principal-stratum strategy.

In the literature of sample size calculation (e.g., [3, 9]), a variety of methods have been proposed for sample size calculation, from which we can select some method to align with the first four of the five attributes of the estimand (i.e., treatment, population, outcome variable, and a population-level summary). Hence, we need to generalize these methods to align with the fifth attribute of the estimand (i.e., how to deal with ICEs).

Fang and Jin (2021) proposed five approaches for generalizing the existing methods of sample size calculation to align with how the ICEs are handled, with each approach aligned with one of the five ICH E9(R1)'s strategies [5]. However, they focused on clinical trials with quantitative and binary outcomes. In this paper, we extend their five approaches to clinical trials with time-to-event (TTE) outcomes. We refer to the "event" in the definition of the TTE outcome as the "primary event", to distinguish it from the censoring event and ICEs.

The rest of the paper is organized as follows. In Section 2, we review one simple method of sample size calculation for clinical trials with TTE outcomes, when there are no ICEs besides censoring. In Section 3, we discuss the application of the five ICH E9(R1)'s strategies to handle ICEs in clinical trials with TTE outcomes. In Section 4, we propose five approaches of sample size calculation when planning clinical trials with TTE outcomes and ICEs, with each subsection devoted to one of the five strategies. We conclude the paper with some discussion in Section 5.

2. CLINICAL TRIALS WITHOUT INTERCURRENT EVENTS

Consider a randomized controlled clinical trial (RCT). Let $Z = 1$ and 0 denote the random assignment to the treatment arm and the control arm (e.g., placebo or standard of care), respectively. Let $v = 0, 1, \cdots, V$ indicate baseline, follow-up visit $1, \cdots$, follow-up visit V, respectively. Setting $t_0 = 0$ as the starting time at baseline, let t_v be the time at visit v, for $v = 1, \cdots, V$. Let A_v be the treatment to be

*Corresponding author.

taken by the subject between visit $v-1$ and v, $v = 1, \cdots, V$. Under the ideal scenario where there are no ICEs, $A_v = Z$, for $v = 1, \cdots, V$.

Let T be the primary TTE outcome variable. For example, in survival analysis, T is the time to death due to a given disease, where "death due to a given disease" is the primary event. Since the ICH E9(R1)'s central questions are in terms of potential outcome [8], let $T^{\boldsymbol{a}}$ denote the potential outcome had the the subject taken a given treatment regime $\boldsymbol{a} = (a_1, \cdots, a_V)$. In particular, we are interested in two potential outcomes, $T^{\boldsymbol{a=1}}$ associated with $\mathbf{1} = (1, \cdots, 1)$ and $T^{\boldsymbol{a=0}}$ associated with $\mathbf{0} = (0, \cdots, 0)$.

Instead of proposing new methods for sample size calculation, in this paper we demonstrate the alignment of the existing methods with how the ICEs are handled. For this aim, consider the simplest setting that assumes exponential distribution on the TTE outcome; that is, $T^{\boldsymbol{a=1}} \sim \exp(\lambda^{\mathbf{1}})$ and $T^{\boldsymbol{a=0}} \sim \exp(\lambda^{\mathbf{0}})$. Hence the average treatment effect (ATE) can be defined in terms of hazard ratio (HR),

$$HR_{ATE} = \lambda^{\mathbf{1}}/\lambda^{\mathbf{0}}. \tag{1}$$

As an estimand of interest, HR_{ATE} has all the first four attributes of an estimand: comparing active treatment regime $\mathbf{1}$ vs. control treatment regime $\mathbf{0}$, a population to be defined via inclusion/exclusion criteria, the TTE outcome variable T, and hazard ratio as a population-level summary. If there are no potential ICEs for every subject, the sample size calculation is straightforward. For example, to achieve $1-\beta$ power under significance level α in a two-sided inequality test, we consider the following formula proposed in [12] for calculating the needed number of events, denoted as E, in the two arms combined,

$$E = 2 \times \left[2 \left(\frac{Z_{1-\alpha/2} + Z_{1-\beta}}{\log HR} \right)^2 \right], \tag{2}$$

where HR is the expected hazard ratio between two arms and Z_a is the upper $100 \times a\%$ quantile of standard normal distribution $N(0,1)$. Note that formula (2) can be extended to any randomization ratio other than 1:1.

To estimate the needed sample size, we need to estimate the survival rates at t_V in the two arms respectively, denoted as $\pi_j = P(T^{\boldsymbol{a=j}} > t_V), j = 1, 0$. Hence, intermediately, as in [12], the needed sample size in the two arms combined is estimated as

$$N' = \frac{E}{1 - (\pi_1 + \pi_0)/2}. \tag{3}$$

A proportion of subjects will be censored due to administrative reasons (e.g., study termination). As in [12], if such proportion is estimated as w, the final needed sample size can be adjusted as

$$N = \frac{N'}{1 - w}. \tag{4}$$

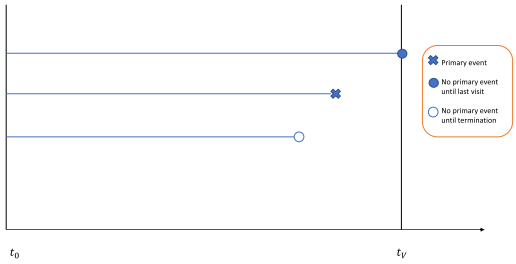

Figure 1. Some examples of observations: one patient with the primary event occurring during the study; one patient having no primary event throughout the study; and one patient censored by the study termination.

Figure 1 presents three examples: one patient with the primary event occurring during the study (the number of such events is denoted by E); one patient having no primary event throughout the study (the proportion of such events in each arm is denoted as π_j); and one patient censored by the study termination (the proportion of such events in two arms combined is denoted as w).

Example 0: Assume that $HR_{ATE} = \lambda^{\mathbf{1}}/\lambda^{\mathbf{0}} = 0.5$, and assume that the survival rate at the last follow-up visit in the control arm would be $P(T^{\boldsymbol{a=0}} > t_V) = 0.6$. This implies the survival rate at the last follow-up visit in the active treatment arm would be $P(T^{\boldsymbol{a=1}} > t_V) = \exp\{0.5 \log(0.6)\} = 0.775$. Also assume that approximately 15% of subjects would be lost to follow-up due to the study termination. To achieve $1 - \beta = 80\%$ power under significance level $\alpha = 5\%$, under the ideal scenario where there are no ICEs other than loss of follow-up, by (2), we can calculate that the needed number of events in two arms combined is $E = 66$. Furthermore, by (3) and (4), we estimate the final needed sample size as

$$N = \frac{66}{[1 - (0.6 + 0.775)/2](1 - 0.15)} = 250, \tag{5}$$

where the first term in the denominator is the averaged event rate at the last follow-up visit, and the second term is to adjust for the loss of follow-up.

Here is a remark on the round-up rule for sample size N. From (5), $N = 248.4706$, which is rounded up to the nearest even number $N = 250$, such that it can be divided into two arms with each arm of 125 subjects.

3. STRATEGIES FOR HANDLING INTERCURRENT EVENTS

In this section, we review the five strategies in ICH E9(R1) for handling ICEs. ICH E9(R1) provided many examples of applying the five strategies to clinical trials with continuous or binary outcomes, but lacked discussion for clinical trials with TTE outcomes. This section itself is novel to fill this gap.

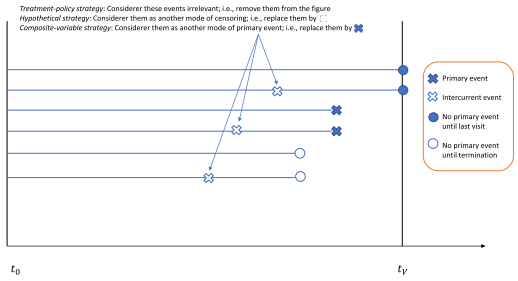

Figure 2. Illustration of the treatment-policy strategy, the hypothetical strategy, and the composite-variable strategy for handling ICEs in clinical trials with TTE outcomes.

We first illustrate the treatment-policy strategy, the hypothetical strategy, and the compositive-variable strategy in Figure 2, where the three subjects in Figure 1 are duplicated and an ICE is added to each duplicated subject.

According to ICH E9(R1), by the treatment-policy strategy, "the occurrence of the intercurrent event is considered irrelevant in defining the treatment effect of interest: the value for the variable of interest is used regardless of whether or not the intercurrent event occurs." To apply the treatment-policy strategy, the data should be collected as usual after the occurrence of an ICE. This is illustrated in Figure 2, where the occurrence of the three ICEs is "ingored" and the value for the outcome is used regardless of whether or not the ICE occurs. Therefore, the treatment-policy strategy can only be applied to handle non-terminal ICEs such as discontinuing treatment, switching treatment, or using rescue medication. Terminal ICEs such as death due to other reasons should be handled by other strategies.

The hypothetical strategy can be applied to handle both non-terminal (e.g., treatment discontinuation) and terminal ICEs (e.g., death due to other reasons). According to ICH E9(R1), by the hypothetical strategy, we need to envisage a scenario "in which the intercurrent event would not occur: the value of the variable to reflect the clinical question of interest is the value which the variable would have taken in the hypothetical scenario defined". This implies that the value that the outcome variable would have taken in the hypothetical scenario is the potential outcome (a.k.a., counterfactual outcome, because it is not observed actually). This is illustrated in Figure 2, where the three ICEs indicated by open checks are considered as another mode of censoring. Therefore, there are two reasons leading to censoring: one is the ICE occurrence handled by the hypothetical strategy and the other is the administrative censoring. In the literature of survival analysis, the administrative censoring is implicitly handled by the hypothetical strategy [4], in which the value of the outcome variable is that would have taken if the subject had continued the assigned treatment after the study termination.

The composite-variable strategy can also be applied to handle both non-terminal ICES (e.g., drop-out due to lack of efficacy) and terminal ICEs. According to ICH E9(R1), by the composite-variable strategy, "an intercurrent event is considered in itself to be informative about the patient's outcome and is therefore incorporated into the definition of the variable". Therefore, we need to re-define the TTE outcome variable to be used in the construction of the estimand of interest. That is, we consider the occurrence of an ICE as another mode of the primary event, as Figure 2 illustrates. For example, progression free survival is defined as the time from the date of randomization to the date of the first documented progression or death.

Now consider the fourth strategy, the while-on-treatment strategy, which has rarely been applied to clinical trials with TTE outcomes. According to ICH E9(R1), by the while-on-treatment strategy, "response to treatment prior to the occurrence of the intercurrent event is of interest". "If a variable is measured repeatedly, its values up to the time of the intercurrent event may be considered relevant for the clinical question". However, an TTE outcome cannot be measured immediately at the time of the occurrence of a non-terminal ICE, and a terminal ICE will preclude the occurrence of the primary event as well. Therefore, in ICH E9(R1), the while-on-treatment strategy is mainly discussed for quantitative and categorical outcomes, but not for TTE outcomes. To order to apply the while-on-treatment strategy to TTE outcomes, we need to envisage a hypothetical scenario where, after the occurrence of a non-terminal ICE, the subjects would stop taking the initially assigned treatments. Like in applying the hypothetical strategy, the while-on-treatment strategy also consider the occurrence of an ICE as another mode of censoring.

Lastly, consider the principal-stratum strategy. According to ICH E9(R1), by the principal-stratum strategy, "the target population might be taken to be the principal stratum in which an intercurrent event would occur. Alternatively, the target population might be taken to be the principal stratum in which an intercurrent event would not occur. The clinical question of interest relates to the treatment effect only within the principal stratum." For TTE outcomes, the principal-stratum strategy is often applied to handle terminal ICEs, and usually the target population is taken to be principal stratum in which an ICE would not occur. Within this principal stratum, there are no ICEs, and therefore the situation becomes the one discussed in Section 2.

Besides the above five strategies discussed in ICH E9(R1), another strategy—to be referred as the competing-risk strategy [1, 10, 13]—can also be applied to handle terminal ICEs in clinical trials with TTE outcomes. Consequently, with the competing-risk strategy in our tool-box, we may apply the aforementioned hypothetical strategy to handle only non-terminal ICEs rather than terminal ICEs. But this is out of scope for this paper.

4. CLINICAL TRIALS WITH INTERCURRENT EVENTS

In planning a clinical trial with TTE outcome, besides the adminstrative censoring that has been adjusted in (4), there may be other types of ICEs. In the following five subsections, we consider five approaches for sample size calculation, with each approach aligned with each of the five ICH E9(R1)'s strategies.

In Section 2, the method of [12] implicitly assumes that the time to censoring is independent of the TTE outcome to simplify the sample size calculation; otherwise, we need to guess or estimate another parameter or more for the association between the censoring and the primary event. This simplification is appropriate because the independence assumption leads to a conservative estimate of the needed sample size. Following this tradition, hereafter, we assume that the time to ICE is independent of the TTE outcome.

4.1 Treatment-policy strategy

Assume that there is one type of ICEs (e.g., treatment discontinuation) and we want to handle it using the treatment-policy strategy. Or, more generally, assume that there is one class of several types of ICEs (e.g., treatment discontinuation, rescue medication) and we want to handle them using the treatment-policy strategy. In practice, we should explicitly specify this strategy in the definition of the estimand, and then align the sample size calculation with it.

By the treatment-policy strategy, equivalently, we conceptually modify the treatment attribute of the estimand: comparing $Z = 1$ vs. $Z = 0$ instead of comparing $\boldsymbol{a} = \mathbf{1}$ vs. $\boldsymbol{a} = \mathbf{0}$. Hence, the new estimand is the intent-to-treat (ITT) effect. Assume that the proportion of subjects who are expected to have an ICE in arm $Z = j$ is r_j, $j = 1, 0$. Assume the hazard in the control arm is the same regardless of whether an ICE occurs; that is, $(T|Z = 0) \sim \exp(\lambda^{\mathbf{0}})$. Meanwhile, to be conservative, assume the hazard in the treatment arm jumps from $\lambda^{\mathbf{1}}$ to $\lambda^{\mathbf{0}}$ after an ICE occurs. This assumption is similar to the jump-to-reference (J2R) assumption in the literature of missing data [2, 11]. Under the J2R assumption, the hazard function in the treatment arm is not a constant, but its expectation is bounded by $(1 - r_1)\lambda^{\mathbf{1}} + r_1\lambda^{\mathbf{0}}$, if—without loss of generality—assuming $\lambda^{\mathbf{1}} < \lambda^{\mathbf{0}}$. For the purpose of sample size calculation, we assume the hazard in the treatment arm is approximated by its bound $(1-r_1)\lambda^{\mathbf{1}}+r_1\lambda^{\mathbf{0}}$; that is, we conservatively assume $(T|Z = 1) \sim \exp\{(1 - r_1)\lambda^{\mathbf{1}} + r_1\lambda^{\mathbf{0}}\}$ for simplicity. Therefore, under these assumptions, the hazard ratio between the two arms, $Z = 1$ vs. $Z = 0$, is diluted from HR_{ATE} to

$$HR_{ITT} = \frac{(1 - r_1)\lambda^{\mathbf{1}} + r_1\lambda^{\mathbf{0}}}{\lambda^{\mathbf{0}}} = (1 - r_1)HR_{ATE} + r_1. \tag{6}$$

Example 1: We continue the discussion of Example 0. Furthermore, assume the proportion of subjects who are expected to have an ICE in each arm is $r_1 = r_2 = r = 10\%$. Using (6) we obtain $HR_{ITT} = (1-0.1)(0.5)+0.1 = 0.55$, and calculate the needed event number as $E = 88$. The survival rate at the Vth visit in arm $Z = 0$ and 1 would be approximately 0.6 and $P(T > t_V|Z = 1) = \exp\{0.55\log(0.6)\} = 0.755$, respectively. Hence the needed sample size is $N = 88/\{[1 - (0.6 + 0.755)/2](1 - 0.15)\} = 322$.

4.2 Hypothetical strategy

Assume that we want to deal with a class of ICEs using the hypothetical strategy. By the hypothetical strategy, we need to envisage a scenario "in which the intercurrent event would not occur". In the hypothetical scenario, $A_v = Z$, for $v = 1, \cdots V$. Hence, we are interested in comparing two treatment regimes, $\boldsymbol{a} = \mathbf{1}$ vs. $\boldsymbol{a} = \mathbf{0}$, and the estimand is HR_{ATE} defined in (1).

Assume that the proportion of subjects who are expected to have an ICE in arm $Z = j$ is r_j, $j = 1, 0$. To calculate the sample size, we first calculate the number of events using (2) with $HR = HR_{ATE}$, calculate N' using (3), and then calculate the sample size N using (4) with w replaced by

$$w^* = w + (r_1 + r_2)/2, \tag{7}$$

which is the sum of the proportion of administrative censoring and the average of the proportions of ICE occurrence in two arms.

Example 2: We continue the discussion of Example 0. Furthermore, we assume the proportion of subjects who are expected to have an ICE in each arm is $r_1 = r_2 = r = 10\%$. We use (2) with $HR = HR_{ATE}$ to calculate the needed event number as $E = 66$. Like in Example 0, the survival rates at the at the last follow-up visit in two regimes $\boldsymbol{a} = \mathbf{0}$ vs. $\boldsymbol{a} = \mathbf{1}$ would be approximately 0.6 and 0.775, respectively. By the hypothetical strategy, we use $w^* = w + r = 0.15 + 0.1 = 0.25$. Hence, the needed sample size is $N = 66/\{[1 - (0.6 + 0.775)/2](1 - 0.25)\} = 282$.

4.3 Composite-variable strategy

Assume that we want to deal with a class of ICEs using the composite-variable strategy. By the composite-variable strategy, we need to re-define the outcome variable to be used in the construction of the estimand of interest. For the TTE outcome, we consider the occurrence of an ICE as another mode of the event of interest. To this aim, let T_{ICE} be the time to the ICE occurrence, and define the following new outcome variable,

$$T^* = \min\{T, T_{ICE}\}. \tag{8}$$

Denote the hazards of T^* in the two arms, respectively, as λ_i^*, $i = 1, 0$. Assume that the distribution of T_{ICE} in arm $Z = i$ is also exponential with hazard ν_i; that is, $(T_{ICE}|Z = i) \sim \exp(\nu_i)$, $i = 1, 0$. Hence, under the assumption that T_{ICE} and T are independent, we have $P(T^* > t|Z = i) = P(T > t|Z = i)P(T_{ICE} > t|Z = i) = \exp(-\lambda^{\boldsymbol{i}}t)\exp(-\nu_i t) = \exp\{-(\lambda^{\boldsymbol{i}} + \nu_i)t\}$ for any $t > 0$,

implying that the distribution of T^* in arm $Z = i$ is exponential distribution with hazard $\lambda_i^* = \lambda^{\mathbf{i}} + \nu_i$, $i = 1, 0$. Note that $P(T_{ICE} > t_V | Z = i) = 1 - r_i$, according to the definition of r_i. Therefore, the hazard ratio associated with T^* between the two arms is

$$HR^*_{ATE} = \frac{\lambda^{\mathbf{1}} + \nu_1}{\lambda^{\mathbf{0}} + \nu_0} = \frac{\log[\exp(-\lambda^{\mathbf{1}} t_V) \times (1 - r_1)]}{\log[\exp(-\lambda^{\mathbf{0}} t_V) \times (1 - r_0)]}. \quad (9)$$

Example 3: We continue the discussion of Example 0. Furthermore, we assume the proportion of subjects who are expected to have an ICE in each arm is $r_1 = r_2 = r = 10\%$. Recall that in Example 0, $\exp(-\lambda^{\mathbf{1}} t_V) = 0.6$ and $\exp(-\lambda^{\mathbf{0}} t_V) = 0.775$. By (9), the adjusted hazard ratio is

$$HR^*_{ATE} = \log[0.775(1 - 0.1)]/\log[0.6(1 - 0.1)] = 0.58.$$

Using (2) with $HR = HR^*_{ATE}$, the needed event number is $E = 106$. Moreover, $P(T^* > t_V)$ in two arms are expected to be $0.6(1 - 0.1)$ and $0.775(1 - 0.1)$, respectively. Hence, the needed sample size is

$$N = \frac{106}{\{1 - [0.6(0.9) + 0.775(0.9)]/2\}(1 - 0.15)} = 328.$$

4.4 While-on-treatment strategy

As discussed in Section 3, the while-on-treatment strategy has rarely been applied to handle ICEs in clinical trials with TTE outcomes, but for completeness, we provide some brief discussion on how to conduct sample size calculation that is aligned with this strategy. Like the hypothetical strategy, the while-on-treatment strategy also considers the occurrence of an ICE as another mode of censoring. The only difference between these two strategies is that the hypothetical strategy assumes that the subjects continue their assigned treatments after the ICE occurrence, while the while-on-treatment strategy assumes that the subjects stop their assigned treatments after the ICE occurrence. Since both strategies consider the occurrence of an ICE as another mode of censoring, the method for sample size calculation is the same.

Example 4: We continue the discussion of Example 0. Furthermore, we assume the proportion of subjects who are expected to have an ICE in each arm is $r_1 = r_2 = r = 10\%$. By the while-on-treatment strategy, the sample size calculation is the same as the one in Example 2.

4.5 Principal-stratum strategy

Assume that we want to deal with a class of ICEs using the principal-stratum strategy. Without loss of generality, consider the target population taken to be principal stratum in which no ICE would occur.

Let C^z be the potential compliance status for $z = 1, 0$. That is, $C^j = 0$ means that an ICE would not occur had the subject been assigned to $Z = j$, $j = 1, 0$. Therefore, the principal stratum in which an ICE would not occur consists of subjects with $C^1 = 0$ and $C^0 = 0$ [6]. Assume that the proportion of subjects who are expected to have ICEs in arm $Z = j$ is r_j, $j = 1, 0$. This implies that

$$\begin{aligned} P(C^1 = 0 \text{ and } C^0 = 0) &= 1 - P(C^1 = 1 \text{ or } C^0 = 1) \\ &\geq 1 - [P(C^1 = 1) + P(C^0 = 1)] = 1 - (r_1 + r_0). \end{aligned}$$

Let N be the needed sample size calculated in Section 2 using (2)-(4). To deal with ICEs using the principal-stratum strategy and considering the principal stratum $\{C^1 = 0, C^0 = 0\}$, to be conservative, we adjust the sample size as

$$N^* = N/(1 - r_1 - r_0). \quad (10)$$

Example 5: We continue the discussion of Example 0. Furthermore, we assume the proportion of subjects who are expected to have an ICE in each arm is $r_1 = r_2 = r = 10\%$. By the principal-stratum strategy considering the principal stratum $\{C^1 = 0, C^0 = 0\}$, the needed sample size is $N^* = N/(1 - r_1 - r_0) = 250/0.8 = 314$.

5. DISCUSSION

ICH E9(R1) emphasized the correct order for conducting clinical trials: "having clarity in the trial objectives and accounting explicitly for intercurrent events when describing the treatment effect of interest at the planning stage should inform choices about trial design, data collection and statistical analysis." Sample size calculation is a key step in study planning and design and it should be aligned with how the ICEs are to be dealt with.

In this paper, we discuss five basic approaches to sample size calculation when planning and designing clinical trials with TTE outcome variables and ICEs, with each approach corresponding to each of those five strategies. In practice, it is not uncommon to consider a combination of several strategies to deal with a combination of several types of ICEs. To be conservative, assume no subject will have more than one ICEs. Hence, there are ten combinations of two different strategies, ten combinations of three different strategies, and so on. We can use a staged procedure to conduct sample size calculation if a combination of several strategies is considered. For example, if we consider a combination of two strategies numbered as 1 and 2, we start with the needed sample size N_0 for the ideal scenario without any ICE, then we adjust the sample size as N_1 with consideration of strategy 1, and then we further adjust the sample size as N_2 with consideration of strategy 2.

To implement any proposed method, we follow the following four steps. First, we categorize all the potential ICEs, along with the corresponding strategies for dealing with them. Second, we explicitly define an estimand of interest that reflects the research objective. Third, together with other investigators, we obtain an expected value of the given effect size and expected proportions of all types of potential

ICEs, along with references that support the expectations. Fourth, we select an appropriate method that is aligned the estimand and calculate the needed sample size.

Received 25 July 2022

REFERENCES

[1] Austin, P. C., Lee, D. S., and Fine, J. P. (2016). Introduction to the analysis of survival data in the presence of competing risks. *Circulation*, 133(6):601–609.

[2] Carpenter, J. R., Roger, J. H., and Kenward, M. G. (2013). Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. *Journal of Biopharmaceutical Statistics*, 23(6):1352–1371. MR3196115

[3] Chow, S.-C., Shao, J., Wang, H., and Lokhnygina, Y. (2017). *Sample size calculations in clinical research*. Chapman and Hall/CRC. MR3753624

[4] Diao, G., Liu, G. F., Zeng, D., Zhang, Y., Golm, G., Heyse, J. F., and Ibrahim, J. G. (2022). Efficient multiple imputation for sensitivity analysis of recurrent events data with informative censoring. *Statistics in biopharmaceutical research*, 14(2):153–161.

[5] Fang, Y. and Jin, M. (2021). Sample size calculation when planning clinical trials with intercurrent events. *Therapeutic Innovation & Regulatory Science*, 55:779–785.

[6] Frangakis, C. E. and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, 58(1):21–29. MR1891039

[7] ICH E9(R1) (2020). ICH E9(R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials.

[8] Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press. MR3309951

[9] Julious, S. A. (2009). *Sample sizes for clinical trials*. CRC Press.

[10] Kalbfleisch, J. D. and Prentice, R. L. (2011). *The statistical analysis of failure time data*, volume 360. John Wiley & Sons. MR0570114

[11] Liu, G. F. and Pang, L. (2016). On analysis of longitudinal clinical trials with missing data using reference-based imputation. *Journal of Biopharmaceutical Statistics*, 26(5):924–936.

[12] Machin, D., Campbell, M., Fayers, P., and Pinol, A. (1997). Statistical tables for the design of clinical studies.

[13] Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, pages 541–554. MR0570114

Yixin Fang
1 North Waukegan Road
North Chicago
IL 60064
USA
E-mail address: yixin.fang@abbvie.com

Man Jin
1 North Waukegan Road
North Chicago
IL 60064
USA
E-mail address: manmandy.jin@abbvie.com

Chengqing Wu
55 Challenger Rd STE 501
Ridgefield Park
NJ 07660
USA
E-mail address: alan1.wu@beigene.com