

Wednesday Challenge (February 2022)

Federico Bonofiglio and Alexander Schacht

06/02/2022

This month's challenge

A phase III cardiovascular superiority study is planned with the following assumptions (simplified, no lost-to-FU, no treatment discontinuation, no non-CV death):

- Hazard rate of active comparator 5% per year,
- Hazard ratio 0.775,
- Accrual time 24 months, total duration 30 months.

Statistical settings:

Using the Logrank test for superiority in time to first MACE event (CV death, MI, stroke) with Alpha 5%, Power 90%, a sample size of about 9900 patients would be needed aiming for 650 patients experiencing a MACE event to show superiority of the new treatment over the comparator in reducing MACE events.

Challenge:

How can uncertainty be displayed with regard to total study duration (TSD) if the assumptions are not exactly met, i.e. comparator hazard rate between 4% and 6% (and hazard ratio between 0.75 and 0.8)?

Background

George and Desu (1974); Rubinstein, Gail, and Santner (1981) suggest that we could find TSD from a survival sample size formula dividing study duration into accrual and follow-up times (exponential model) by means of numerical methods. George and Desu focus on the problem of finding the optimal accrual time in studies with no censoring and no additional follow-up time. Rubinstein et al. put an additional focus on censoring and follow-up time, stressing the robustness of the exponential model with regard to the non-parametric log-rank test. The main contribution of these authors is with regard to study planning, for instance, by providing better justifications for allocating time between an accrual and follow-up phase such to optimize the requested number of failures and overall sample size. Therefore, optimization with regard to TSD is not the overt goal and no standard approach to analyze this problem seems yet available (Machida, Fujii, and Sozu 2021).

Machida et al. propose a method to analyze the probability distribution of TSD by keeping accrual time fixed, thus focusing on follow-up duration by varying sample size and underlying hazards. Their method however does not work for big sample size. Finally, the literature on group sequential design might also be worth being explored, but we do not give any reference here.

Method used

We want to analyze how TSD changes by varying hazard ratio (HR) and baseline hazard (λ_c), using an exponential model allowing for fixed sample size, accrual time (a), and type I and II errors. Otherwise said we want to answer the question ‘how does follow-up time changes provided we keep resources for sample size, accrual time, and power fixed, but let the assumed baseline hazard and HR vary?’

We use `nSurvival` from package `gsDesign` to reproduce the reference example ($\lambda_c = 0.05$, $HR = 0.775$, $a = 24$, $TSD = 30$, $\alpha = 0.05$, $\beta = 0.1$, two-sided test) obtaining a number of requested failure equal to 651 and a sample size of 10157 which is more conservative than the 9900 above. For the purpose of our exercise we stick to this approximation.

Next, we optimize `nSurvival`, denoted as the function $N(\tau|\lambda_c, HR)$ of unknown TSD, τ , given λ_c and HR, while keeping I-II error and accrual time fixed. This can be rewritten into a minimization problem of the following loss function,

$$||N(\tau|\lambda_c, HR) - N_{reference}||,$$

which has a minimum at 0, under the constraint $\tau \geq a$, and where $N_{reference}$ is the sample size equal to 10157 obtained under the reference settings.

We use `optimize` with numerical derivation to minimize the above loss function along pre-specified sequences of baseline-hazard and HR values, and using a lower bound of $\tau = a + 1 = 25$, or a minimum follow-up time of one month. We recalculate the requested number of failures by varying HR.

Code is available upon request.

Results

The animated figure below plots TSD against HR for varying values of the baseline hazard (hazard of active comparator). Points increase in size according to the number of required failures. The green lines intersect the reference result. The static blue points are calculated under the reference baseline hazard of 0.05.

The plot shows that both required number of failures and TSD increase for increasing HR (decreasing treatment effect). Moreover, TSD generally decreases for increasing baseline hazard.

The optimal TSD is found for increasing treatment effect ($HR \rightarrow 0$) and for increasing baseline hazard, which also yields the smallest number of required failures. Eventually, as the hazard in the active comparator becomes so high, TSD reaches its lower bound for all HR values.

For PDF output please refer to the separate `Wednesday_challenge.gif` file.

Discussion

The lower bound of 1 month follow-up was arbitrarily chosen under satisfaction of the constraint. The theoretical minimum is $TSD = \text{accrual time}$ (George and Desu 1974), i.e. when no follow-up time occurs. In such case a conservative approximation is $TSD \approx d/m$, where d is the total required failures in the study and m is the failures accrual rate. This means that in our setting, as TSD approaches its lower bound, the rate of failures m must increase in a shorter time range for a given fixed sample size and for decreasing treatment effect in order to maintain the wished power. Whether this is realistic it depends on further assumptions, including on m . Therefore, a method to more comprehensively assess further assumption could be to extend our optimization task to several parameters simultaneously, e.g. to both accrual time and follow-up time, and to better assess how power changes accordingly.

Bibliography

- George, Stephen L., and M. M. Desu. 1974. “Planning the Size and Duration of a Clinical Trial Studying the Time to Some Critical Event.” *Journal of Chronic Diseases* 27 (1-2): 15–24. [https://doi.org/10.1016/0021-9681\(74\)90004-6](https://doi.org/10.1016/0021-9681(74)90004-6).
- Machida, Ryunosuke, Yosuke Fujii, and Takashi Sozu. 2021. “Predicting Study Duration in Clinical Trials with a Time-to-Event Endpoint.” *Statistics in Medicine* 40 (10): 2413–21. <https://doi.org/10.1002/sim.8911>.
- Rubinstein, Lawrence V., Mitchell H. Gail, and Thomas J. Santner. 1981. “Planning the Duration of a Comparative Clinical Trial with Loss to Follow-up and a Period of Continued Observation.” *Journal of Chronic Diseases* 34 (9-10): 469–79. [https://doi.org/10.1016/0021-9681\(81\)90007-2](https://doi.org/10.1016/0021-9681(81)90007-2).