Advances in Time-to-Event Analyses in Clinical Trials
University of Pennsylvania, 2023

Panelist: Mei-Cheng Wang, Department of Biostatistics, Johns Hopkins University
Congratulations to Devan, Lu, Zhenzhen and Fan!

This discussion:

- will focus on use of multiple endpoints in clinical trials.
- Comments mainly related to presentations of Devan Mehrotra and Lu Tian.
Lu Tian’s work

Analyzing duration of response (DOR)

$Y_1$: time to response; $Y_2$: DOR
If $Y_1 < T$ then $T = Y_1 + Y_2$. If $Y_1 \geq T$ then $Y_1 = \infty$.
$T$: time to endpoint
$C$: censoring time

- **Induced informative censoring:**
  The censoring time for $Y_2$ is $\max\{C - Y_1, 0\}$. If $Y_1$ is correlated with $Y_2$ then the observation of $Y_2$ is subject to informative censoring.

- **Standard Kaplan-Meier estimation is biased.** Lu Tian presented bias-corrected estimation of $S_{Y_2}(t)$ (same as Lin-Ying estimator.).

- **Huang and Louis (1998, Biometrika):** Joint estimation of a survival function and mark variables.
  - Mark variable could be DOR, medical cost, biomarker measurement, etc.
Devan Mehrotra’s work

Analyzing composite or multiple endpoints

$Y_j$: the jth gap time
$T_j = Y_1 + ... + Y_j$: time to jth endpoint
$t$: time since 0

Data example: CV death, Non-fatal stroke, Non-fatal MI, Coronary revasc≥30 days after randomization, Unstable angina

- Prentice, Williams & Peterson model (1981) assume the jth endpoint occurs with the hazard

$$
\lambda_j(t - t_{j-1} \mid X, N^H(t^-)) = \lambda_{0j}(t - t_{j-1}) \exp\{\beta'_j X + \phi(N^H(t^-))\}
$$

$\lambda_{0j}(\cdot)$: baseline hazard function. $N^H(t^-)$: history of events prior to $t$
Can PWP Model be used in clinical trials?

- For the purpose of testing, need simplification:
  \[ \lambda_j(t - t_{j-1} \mid X, N^H(t^-)) = \lambda_0 j(t - t_{j-1}) \exp\{\beta'_j X + \phi(N^H(t^-))\} \]
  \[ \lambda_j(t - t_{j-1} \mid X, N^H(t^-)) = \lambda_0 j(t - t_{j-1}) \exp\{\beta'_j X\} \]
  - it requires that each gap time carries NO memory from the event history!

- PWP considered ordered endpoints so \( \lambda_j \) has concrete meaning; e.g., HIV - AIDS - Death

  **What does \( \lambda_j \) mean if multiple endpoints have arbitrary order?** (CV death, Non-fatal stroke, Non-fatal MI, etc.)

- Competing risks issues: Since ‘death’ must be part of censoring for other endpoints, what type of hazards is PWP model estimating? - **Cause-specific hazard**

- If PWP model estimates cause-specific hazards, does the hazard provide proper causal interpretation for clinical trials?
Statistical approaches

- **Anderson-Gill approach**: The AG model assumes constant intensity ratio; the multiple event process does not carry memory:
  \[ \lambda(t \mid N^H(t^-), X) = \lambda_0(t) \exp\{\beta' X\} \]
  - AG model is similar to a Poisson process; ‘memoryless’ is a very strong assumption!
  - Does not handle absorbing event such as death.

- **Wei-Lachin, Brown, Kost-McDermott, Stouffer, Lachin-Bebu approaches**: Cox PH model for each component, average resulting HR estimates with differential weights.
  - Conceptually, better than PWP and AG approaches.
  - In the presence of CV-death, the hazard for other multiple events (Non-fatal stroke, Non-fatal MI, etc.) are cause-specific hazards.
  - Are we comfortable with the use of cause-specific hazards to evaluate trial effects?
Statistical approaches

- **Win ratio statistic**: Aggregate pairwise subject-level between-treatment comparison of survival times based on sequential order of endpoint importance (Pocock et al., 2012).
  - The win ratio statistic is simple in its format but hard to understand.
  - Fortunately, in semi-competing risks setting, the aim of hypothesis testing has been clarified to test on the observable part of the hazards of latent variables; see Luo et al. (Biometrics, 2015) and Mao Lu (2021).
  - Analysis results possess causal interpretation!

- **Claggett et al. approach**: Quantify mean cumulative count of events over time using AUC to compare treatments. AUC is an overall summary measure of all the events.
  - Conceptually, this is an appropriate approach in the presence of multiple endpoints (including death).
  - Analysis results possess causal interpretation!
  - Why worst performance in simulation?