Estimand and analysis consideration in a Phase III study of CAR-T (BELINDA) with delayed treatment effect – a case study of lymphoma

Jessie Gu, on behalf of Novartis CTL team

Oncology Full Development Biostatistics, Novartis

BACKGROUND

- Patients with aggressive B-cell non-Hodgkin lymphoma (NHL) can be cured by frontline immunotherapy, however, around 1/3 will experience early relapse or are refractory (r/r).
- These patients have a poor prognosis even if they receive autologous stem cell transplant (HSCT).
- Tisagenlecleucel (Kymriah, CART-19) is a second generation CAR-T cell product that uses autologous peripheral blood T cells that have been genetically modified ex vivo to target CD19 on the surface of B cells.

Figure 1: Tisagenlecleucel chimeric antigen receptor design

BELINDA is a randomized, open label, multicenter phase III trial comparing the efficacy, safety, and tolerability of tisagenlecleucel to standard of care (SOC) in adult patients with relapsed/refractory aggressive B-cell NHL.

METHODS

Figure 2: BELINDA study design

INTERVENTION/TREATMENT

A total of 318 patients will be randomized 1:1 to the two treatment arms:
- Tisagenlecleucel: Tisagenlecleucel after optional bridging and lymphodepleting chemotherapy.
- Standard of Care (SOC): Platinum-based immunotherapy followed in chemotherapy (i.e. R-ICE, R-GDP, R-DHAP, R-GemOx), therefore expect similar treatment effect in both arms prior to tisagenlecleucel infusion

PRIMARY ENDPOINT

Event-free survival (EFS) – time from randomization to the earlier of the following:
- SD/PO by BIRC after 12 week assessment
- Death at any time

NON-PROPORTIONAL HAZARDS

- Patients in the treatment arm receive bridging chemotherapy while waiting for the manufacturing of tisagenlecleucel, i.e. ~6 wks (1.38 months).
- Same immunotherapy regimens for bridging and SOC (i.e. R-ICE, R-GDP, R-DHAP, R-GemOx), therefore expect similar treatment effect in both arms prior to tisagenlecleucel infusion
- Treatment effect starts to manifest after initiation of tisagenlecleucel infusion
- Treatment effects plateauing after >9 months

WEIGHTED LOG-RANK (LR) TEST

- LR test can be viewed as weighted test with weight=1, it is most powerful under proportional hazard assumption
- The optimal weighted LR test under NPH has weights proportional to the log of hazard ratio (λ)

SIMULATION

Table 1: Three-piece exponential EFS curve with a total of 318 patients randomized 1:1 to the two treatment arms – 2 simulated scenarios

Table 2: Power of different weighted LR tests from simulation

CONCLUSIONS

- Weighted LR test and weighted/piecewise HR estimates can be useful in presence of NPH
- Be cautious about the interpretation of the weighted LR test and estimates: weights for FH test or max combo test depends on the underlying empirical survival function
- Interpretation of piecewise LR test is straightforward, however mis-specification of the delayed period (period HR=1) may lead to biased estimation (Ku et al. 2018)
- Use weighted or unweighted LR test as the study primary analysis depends on the objective of the trial:
  - Comparison during all periods after randomization ? (e.g. unweighted LR test)
  - More focus on comparison during period where differences are expected? (e.g. weighted or piecewise test)
- Multiple approaches may be necessary to summarize results

ACKNOWLEDGEMENT

Many thanks for the strong support from Novartis statistical methodology group and the entire CTL019 project team for providing valuable input for the estimand discussion and simulation.

REFERENCE

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