Supporting decisions at the interim analysis for a study in advanced cervical cancer

An application of Bayesian dynamic borrowing survival models informed by historical trial data

MSR142

View all Parexel's posters at ISPOR Europe 2024

Daniel J. Sharpe[1], Meghann Gregg[2], Andrew R. Wilson[2,3], Riccardo Ressa[1], Emtiyaz Chowdhury[1], Tuli De[2], Jackie Vanderpuye-Orgle[2]

[1]Parexel, London, UK
[2]Parexel, Durham, North Carolina, USA
[3]University of Utah, Salt Lake City, Utah, USA

Background

- Extrapolation of survival outcomes from immature trial data without support from external information may lack justification and be unappealing to decision-makers[1]
- KEYNOTE-826 is a phase III study of pembrolizumab plus chemotherapy with or without bevacizumab (PEMBRO+CHEMO +/-BEV, vs CHEMO+/-BEV) in first-line advanced cervical cancer (1L aCC)[2]
  - in an initial appraisal, NICE considered the interim data cut of KEYNOTE-826 (minimum 15 months follow-up)[2] and recommended PEMBRO+CHEMO+/-BEV via managed access[3]
  - NICE later issued a final positive recommendation for routine commissioning following the availability of the final data cut (minimum 30 months follow-up)[4,5]
- Bayesian dynamic borrowing (BDB)[6,7] offers an appealing approach to improve reliability and transparency of treatment effect estimates from immature data by incorporating longer-term external observations
- We designed a BDB model to project overall survival (OS) from the interim data cut of KEYNOTE-826, informed by historical control data from the phase III GOG-240 study[8] of CHEMO+BEV vs CHEMO in 1L aCC (maximum 50 months follow-up)
- We aimed to assess the potential benefits that may have been offered by employing BDB with historical control data to predict the long-term efficacy of PEMBRO+CHEMO+/-BEV at the interim analysis, including considerations related to:
  - decision risk: the BDB model is based on demonstrably conservative assumptions
  - statistical uncertainty: the BDB model directly incorporates external data with extended follow-up
  - decision speed: the BDB model avoids naïve speculation in OS projections, which was a key concern in the initial NICE appraisal, and so may have reduced the required number of meetings

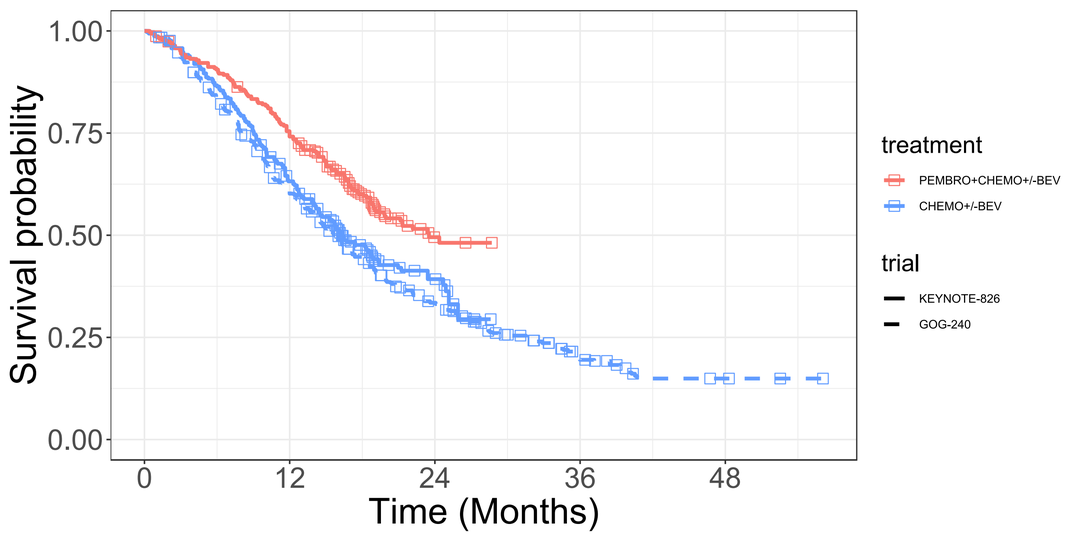
Methods

- Reconstructed patient-level data from KEYNOTE-826 and GOG-240 were used to estimate generalized gamma distributions representing OS in these trials, via BDB
  - in the BDB models for both the experimental and control arms of KEYNOTE-826, the external observations were pooled from the two arms of GOG-240
  - BEV use was a stratification factor in KEYNOTE-826[2] and GOG-240 demonstrated superiority of CHEMO+BEV vs CHEMO[8]. Thus, the GOG-240 observations were reweighted to reflect the proportion of patients treated with CHEMO+BEV in KEYNOTE-826 (namely, 63.0%)
  - the reweighted GOG-240 data are highly consistent with the control arm observations from KEYNOTE-826 (Fig. 1), implying that remaining between-trial differences are minimal
- The BDB models expressed a relatively strong *a priori* preference for KEYNOTE-826 parameter estimates to be similar to concomitantly estimated GOG-240 values
  - this preference was enforced via commensurate prior distributions[6,7], in which the parameter estimates for the external data (GOG-240) model were used as the means in normal prior distributions for the current data (KEYNOTE-826) model
  - the variance parameters of the commensurate priors were drawn from a gamma hyperprior distribution with appreciable probability density at both low and high values, which lead to strong and weak propensity for the model to "borrow" from the external data, respectively
  - in this way, the BDB model automatically "learns" to "take-or-leave" the historical control information (from GOG-240) to supplement predictions for the current study (KEYNOTE-826), based on a user-defined propensity for borrowing, separately for each model parameter
- To assess the effect of the historical control data on survival extrapolations and surrounding uncertainty, a second Bayesian model employing vague prior distributions was fitted to the KEYNOTE-826 data for comparison
- Model estimates were reported as posterior means [with 95% credible intervals]

Results and Discussion

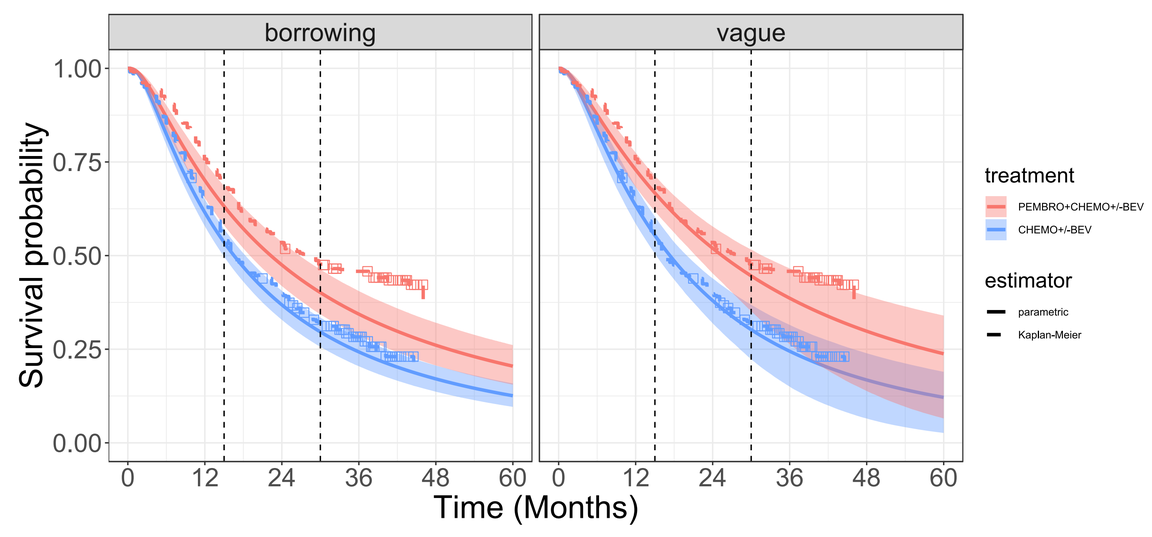
- Leveraging historical trial data for CHEMO+/-BEV led to more conservative and less uncertain estimates of 30-month OS for PEMBRO+CHEMO+/-BEV: 40.1% [34.3-46.3%] BDB vs 44.6% [35.1-56.1%] vague vs 47.4% observed (Fig. 2)
  - the BDB model formulation employed here apparently imposed an overly conservative *a priori* expectation, through the strict treatment waning assumption implicitly invoked by the use of historical control data, that did not reflect the true sustained benefit with the addition of PEMBRO
- Utilizing the GOG-240 observations yielded highly similar and less uncertain estimates of 30-month OS for CHEMO+/-BEV: 29.6% [25.6-33.4%] BDB vs 30.2% [22.3-36.9%] vague vs 31.6% observed
- Thus, the BDB model yielded a smaller point estimate for the treatment effect on 30-month OS probabilities than the vague Bayesian model, but also reduced uncertainty and consequently achieved statistical significance
  - in contrast, the point estimate for the treatment effect from the vague Bayesian model was more accurate, but there were overlapping uncertainty intervals in the longer-term OS predictions for the experimental vs control arms
- In principle, performance of the BDB model for the PEMBRO+CHEMO+/-BEV arm of KEYNOTE-826 could have been improved by using longer-term historical study data for a treatment that exhibits similar durable benefit as this intervention
  - however, no such data were available at the time of the interim analysis of KEYNOTE-826, owing to the lack of previous studies of immunotherapies in 1L aCC[9]
  - while the use of historical control data in Bayesian survival models is inherently limited, such data are highly useful to explore well-defined scenarios to estimate a reasonable lower bound for longer-term treatment effect and avoid naïve speculation in projections
- Further work is required to assess the sensitivity of BDB estimates to the hyperprior distributions for the variance parameters of the commensurate priors
  - the conservativeness of the treatment effect estimate from BDB could be attenuated by tuning hyperprior distributions to reduce the propensity of the PEMBRO+CHEMO+/-BEV arm model to borrow from the historical control data, although this would then induce greater uncertainty
  - moreover, simpler static Bayesian borrowing[6,10] approaches may prove sufficient, especially since this survival model does not include parameters for the effects of baseline covariates

Figure 1: Kaplan-Meier curves for the reconstructed data from the interim analysis of KEYNOTE-826 and the final analysis of GOG-240.



The GOG-240 data have been reweighted to reflect BEV use in KEYNOTE-826. Squares indicate censoring.

Figure 2: Estimates for OS in KEYNOTE-826 from Bayesian models informed by historical trial data for the control arm via dynamic borrowing (BDB) and without prior external data (vague), compared to later observations.



Vertical lines indicate minimum follow-up duration in the interim (to which the parametric models are fitted) and final data cuts. Shaded areas indicate 95% credible intervals. Squares indicate censoring in the reconstructed Kaplan-Meier curves.

Conclusions

- Both BDB and uninformed models successfully forecasted the longer-term superiority of PEMBRO+CHEMO +/-BEV (vs CHEMO+/-BEV) in 1L aCC from KEYNOTE-826 interim data
- BDB enabled the transparent implementation of conservative assumptions surrounding the longer-term efficacy of the intervention and substantially reduced statistical uncertainty in the estimated treatment effect on survival beyond the available follow-up at the interim analysis
  - even though BDB extrapolations for PEMBRO+CHEMO+/-BEV were pessimistic, incorporation of historical control data arguably provided a more compelling demonstration of the superior long-term efficacy of the intervention at the interim
- Payers may consider BDB as a sophisticated method for mitigating decision risk when presented with immature OS data, provided relevant historical study data with extended follow-up are available

REFERENCES

[1] Bullement A, et al. *Med Decis Making* 2023; 43(5): 610-620.
[2] Colombo N, et al. *N Engl J Med*. 2021; 385(20): 1856-1867.
[3] National Institute for Health and Care Excellence. TA885. 2023.
[4] Monk BJ, et al. *J Clin Oncol*. 2023; 41(36): 5505-5511.
[5] National Institute for Health and Care Excellence. TA939. 2023.
[6] Hobbs BP, et al. *Biometrics* 2011; 67(3): 1047-1056.
[7] Hobbs BP, Sargent DJ, and Carlin BP. *Bayesian Anal*. 2012; 7(3): 639-674.
[8] Tewari KS, et al. *Lancet* 2017; 380(10103): 1654-1663.
[9] Martinez-Cannon BA and Colombo I. *Cancer Drug Resist*. 2024; 7(23): 10.20517/cdr.2023.120.
[10] Biard L, et al. *Contemp. Clin. Trials Commun.* 2021; 21: 100709.

Contact: Daniel.Sharpe@parexel.com

www.parexel.com