Extending Bayesian evidence synthesis to include historical trial data for improved survival extrapolations from early data cuts in metastatic non-small cell lung cancer (mNSCLC)

MSR143

Daniel J. Sharpe,[1] Georgia Yates,[1] Llenalia García-Fernández,[2†] Yong Yuan,[3] Adam Lee,[4] Mohammad Ashraf Chaudhary[3†]

[1]Parexel, London, UK; [2]Parexel, Madrid, Spain; [3]Bristol-Myers Squibb, Princeton, NJ, USA; [4]Bristol-Myers Squibb, Uxbridge, UK; [†](at the time the study was conducted)

Background

- The CheckMate 9LA phase 3 trial is evaluating the efficacy of nivolumab plus ipilimumab plus two cycles of chemotherapy (NIVO+IPI+CHEMO) (versus four cycles of CHEMO) in patients with mNSCLC.[1,2]
- The CheckMate 227 Part 1 phase 3 trial has demonstrated durable survival benefit associated with NIVO+IPI, versus CHEMO, in patients with stage IV or recurrent NSCLC.[3]
- The data cuts of CheckMate 227 Part 1 have around 24 months of additional follow-up compared to the data cuts of CheckMate 9LA that were available at the same time.
- These two studies are closely related, namely by:
  - —expected comparability of survival outcomes for patients receiving NIVO+IPI+CHEMO and NIVO+IPI, and between patients receiving different types of CHEMO
  - —similar patient population (e.g., with respect to disease stage and tumor histology) in the first-line setting

  hence, it is desirable to leverage CheckMate 227 Part 1 observations to inform survival extrapolations in CheckMate 9LA, where data are less mature.
- Previous studies have demonstrated that using a Bayesian framework to support survival extrapolations from early data cuts with appropriate historical trial data can produce models that anticipate longer-term phenomena specific to novel classes of therapy[4,5], such as survival plateaus.
- Here, we extend the Bayesian multi-parameter evidence synthesis[6-8] (B-MPES) framework to incorporate historical trial data. In B-MPES, the external data contribute to the model predictions through terms based on one-year conditional survival probabilities at selected timepoints.
- To illustrate the method, we estimate a B-MPES model for an early data cut of CheckMate 9LA with 1 year of minimum follow-up. In this B-MPES model, survival extrapolations are informed by observations from the 3-year data cut of CheckMate 227 Part 1, and are further supported by observations from the 5-year data cuts of CheckMate 017+057[9] (trials of NIVO in second-line advanced NSCLC, comprising patients with squamous and non-squamous histology, respectively).

Methods

- We fitted a B-MPES model based on a 2-knot spline function to the 1-year data cut of CheckMate 9LA, leveraging the following external information sources:
  - —for the period 2-3 years: observations from CheckMate 227 Part 1 (both arms)
  - —for the period 4-5 years: observations from CheckMate 017+057[9] (experimental arm only)
  - —for the period 4-15 years: observations from the Surveillance, Epidemiology, and End Results (SEER) registry dataset[10], for patients diagnosed with advanced or metastatic NSCLC (control arm only)
  - —for the period 23-25 years: trial-matched general population mortality data.[11]
- The control and experimental arm predictions are linked within B-MPES via an imposed hazard ratio condition. Here, to reflect *a priori* belief of durable response in the NIVO+IPI+CHEMO arm of CheckMate 9LA, we specify an expected hazard ratio that tapers linearly from 0.8 to 1 for the period 6 to 10 years and remains at 1 thereafter. In this way, the SEER data indirectly informs the experimental arm estimates, while allowing for the NIVO+IPI+CHEMO hazards to deviate from the SEER data. Although, some expert judgement is required in the selection of these expected values, and these inputs may be highly dependent on the disease area and therapies.
- The incorporation of the second-line NIVO monotherapy data is intended to provide further *a priori* information on the expected emergent survival plateau for patients receiving dual immunotherapy. These data are believed to be more suitable for informing NIVO+IPI+CHEMO survival than the SEER data, since they reflect the possibility of durable survival benefit with immunotherapy, but are an inferior external information source compared to CheckMate 227 for the current application.
- The B-MPES model was fitted using the Stan program.[12]
- The model performance is judged by comparison with:
  - —later observations from a more mature data cut of CheckMate 9LA, with 4 years of minimum follow-up
  - —a standard parametric model (SPM), selected naively based on statistical goodness-of-fit criteria, with adjustment by general population data when the predicted SPM hazards are exceeded by these data.
- A scenario analysis was conducted where CheckMate017+057 are the only sources of historical trial data.

Results

- The best-fitting SPM was a dependent log-logistic distribution. Adjustment by general population data was performed from 15 years onwards.
- The B-MPES model yields significantly more optimistic short-term NIVO+IPI+CHEMO extrapolations than the SPM (Table 1, Fig. 1). The B-MPES model better anticipates the emergent survival plateau that is apparent in the later data cut of CheckMate 9LA, while the SPM underpredicts subsequently observed survival.
- All parametric models underestimate the subsequently observed CHEMO survival (Fig. 1). This feature is probably attributable to patients in the CHEMO arm receiving subsequent immunotherapy, which is not accounted for in any of the models.
- The hazard (Fig. 1) and one-year conditional survival (Fig. 2) curves estimated from B-MPES have a second inflection point at around 6 years.
- Thus, the long-term NIVO+IPI+CHEMO survival predictions (e.g., at 20 years) from B-MPES and the SPM are in close agreement (Table 1).
- Further disparity in the behavior of the two models can be seen in the time-dependence of the NIVO+IPI+CHEMO vs CHEMO hazard ratio (Fig. 3). In the B-MPES model, the hazard ratio tapers more slowly, in accordance with the imposed condition.
- Survival estimates from the B-MPES scenario model are significantly smaller than estimates from the base case B-MPES model (Table 1).

Figure 1. Survival (above) and hazard (below) functions and 95% uncertainty intervals for the B-MPES model and unadjusted SPM fitted to the 1-year data cut of CheckMate 9LA, compared to trial observations from the 4-year data cut.

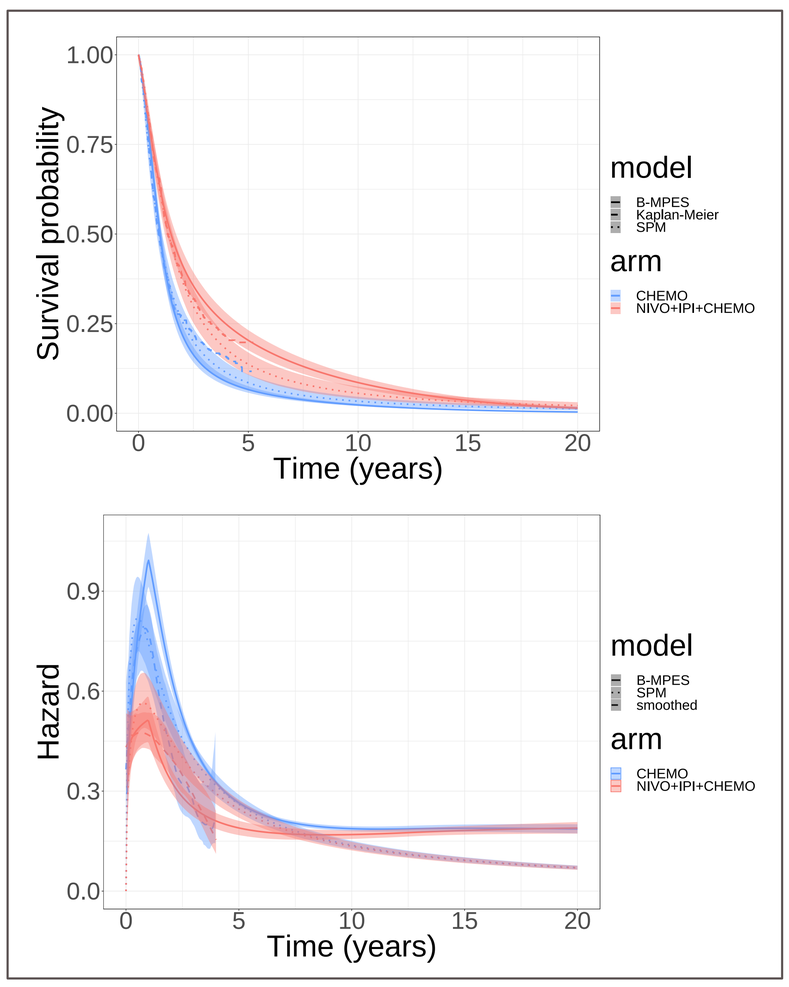


Figure 2. One-year conditional survival probabilities for the B-MPES model and unadjusted SPM fitted to the 1-year data cut of CheckMate 9LA, compared to Kaplan-Meier (KM) estimates from the trial, and probabilities derived from external data sources, including historical trial data (supp).

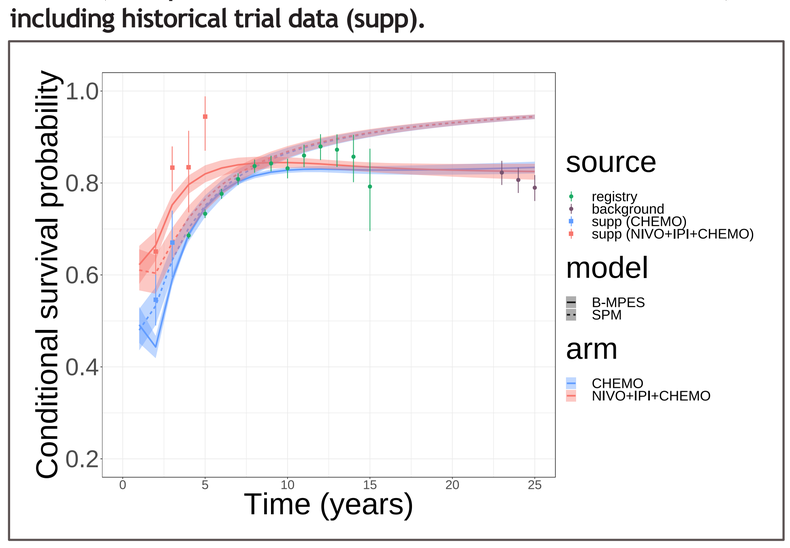
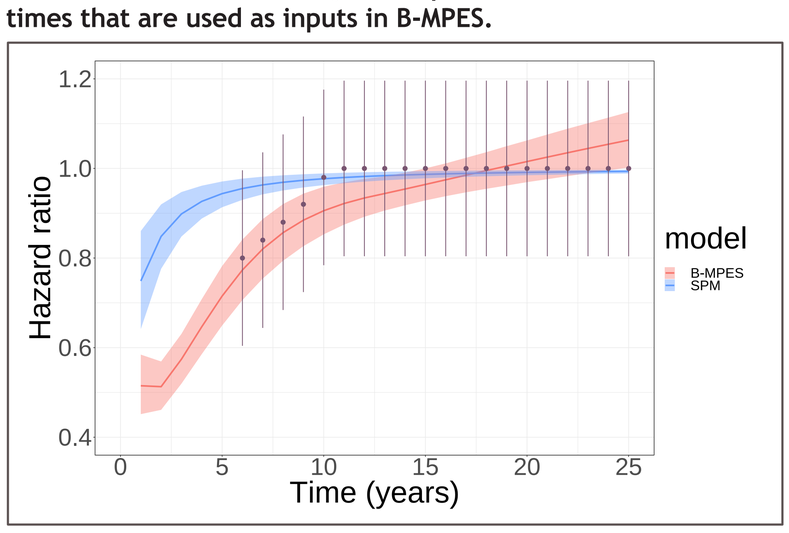


Table 1. Milestone NIVO+IPI+CHEMO survival probability estimates (%) and 95% uncertainty intervals from B-MPES models and an adjusted SPM fitted to the 1-year data cut of CheckMate 9LA, and trial observations from the later data cut.

Time (years)	Survival probability (%) by model			
	B-MPES	B-MPES (scenario)	Adjusted SPM	Kaplan-Meier
4	24.8 [21.7-28.0]	19.3 [16.3-22.5]	17.9 [14.5-21.3]	21.0 [17.2-25.6]
5	20.3 [17.6-23.2]	15.3 [12.7-18.1]	13.7 [10.8-16.6]	-
10	8.6 [7.0-10.2]	5.9 [4.7-7.4]	5.6 [4.0-7.3]	-
20	1.4 [1.1-1.8]	1.0 [0.8-1.3]	1.9 [1.4-2.5]	-

Figure 3. Predicted hazard ratios and 95% uncertainty intervals for the B-MPES model and unadjusted SPM fitted to the 1-year data cut of CheckMate 9LA, and expected values at selected times that are used as inputs in B-MPES.



Discussion

- Bayesian methods provide a formal framework to incorporate historical trial data into a survival model, and are superior to simple post-hoc adjustment, since all information sources are integrated simultaneously, thereby allowing for rigorous uncertainty analysis and the attenuation of overfitting to any one data source.
- The SEER data appear to be a relatively good representation of CHEMO conditional survival on extended timescales but are an overly pessimistic approximation to NIVO+IPI+CHEMO survival. Nonetheless, inclusion of SEER data in the B-MPES model results in the qualitatively correct hazard trend and gives clinically plausible results.
- The historical trial data can have a strong influence on the short-term extrapolations in B-MPES. Intuitively, the use of more conservative data sources can substantially decrease survival estimates. This effect is clear from the scenario analysis, where the only source of historical trial data is from immuno-monotherapy in the second-line setting, which does not fully capture the benefit specific to the dual immunotherapy, and also reflects a patient population with poorer prognosis.
- Even so, imperfect or confounded historical trial data remains useful within Bayesian survival models since such data provides broad information on longer-term phenomena such as durable response, to which the survival model would otherwise be naïve.

Conclusions

- When relevant historical trial data are integrated into a survival model appropriately, this information can improve the reliability of short- and long-term extrapolations and is especially useful when trial data are immature.
- Historical trial data are an especially valuable source of external information since they can be specific to a novel therapy or class thereof. Hence, incorporation of these data can lead to survival models that correctly anticipate treatment-specific effects such as durable response.
- Future work will further explore the impact of using alternative combinations of data sources with B-MPES, and will also involve performing additional sensitivity analyses surrounding the choice of the hazard ratio condition.

References

1. Paz-Ares L, et al. *Lancet Oncol*. 2021; 22(2): 198-211.
2. Paz-Ares L, et al. *J Thorac Oncol*. 2023; 18(2): 204-222.
3. Brahmer JR, et al. *J Clin Oncol*. 2023; 41(6): 1200-1212.
4. Soikkeili F, et al. *Value Health*. 2019; 22(9): 1012-1017.
5. Biard L, et al. *Contemp. Clin. Trials Commun*. 2021; 21: 100709.
6. Guyot P, et al. *Med Decis Making*. 2017; 37(4): 353-366.
7. Chaudhary MA, et al. *Med Decis Making*. 2023; 43(1): 91-109.
8. Sharpe DJ, et al. Poster presentation at ISPOR 2023; May 2023; Boston, USA. Abstract MSR29
9. Borghaei H, et al. *J Clin Oncol*. 2021; 39(7): 723-733.
10. Surveillance Research Program, National Cancer Institute. SEER*Stat. 8.4.0.1. 2022.
11. World Health Organization, WHO methods and data sources for life tables 1990-2019.
12. Carpenter B, et al. *J Stat Softw*. 2017; 76(1): 1-32.

Acknowledgements

This study was supported by Bristol Myers Squibb.

Presented at ISPOR Europe 2023; November 12-15, 2023; Copenhagen, Denmark

Email: Adam.Lee@bms.com