Formulating Bayesian poly-hazard models for informed and clinically interpretable lifetime survival extrapolations in advanced non-small cell lung cancer (aNSCLC)

MSR13

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

- Immuno-oncology therapies, such as nivolumab plus ipilimumab (NIVO+IPI), induce durable response in a proportion of patients with aNSCLC, and hence the long-term survival extrapolations make a substantial contribution to the estimated lifetime benefit of the intervention.[1-3]
- Since there is typically only limited follow-up data available at the time of initial health technology assessments (HTAs), it is strongly advisable to use relevant external data sources to support survival extrapolations.[4,5]
- Bayesian methods provide a holistic and statistically rigorous framework to incorporate external data into parametric survival models, and thus may help to attenuate issues of overfitting that may arise when performing crude post-hoc adjustment.[6]
- However, there is currently limited guidance on best practices for the application of Bayesian survival models in HTAs[7], and these models are substantially more complex than their frequentist analogues. In particular, there are many different formulations of Bayesian survival models that invoke alternative assumptions and could feasibly yield disparate predictions.
- Here, we investigate a specific Bayesian formulation of poly-hazard models[8], which have attractive features of flexibility, straightforward clinical interpretation, and absence of mandatory user-specified auxiliary parameters.
- We apply the approach to overall survival data for NIVO+IPI from the primary data cut of CheckMate 227 Part 1[2], a phase 3 randomized trial in first-line stage IV or recurrent NSCLC, with 29.3 months of minimum follow-up.

Methods

- We consider a poly-hazard model comprising two separate contributions to the observed hazards, which represent "background" and disease-specific components:

$$h(t) = h_{\text{disease}}(t) + h_{\text{background}}(t).$$

- In a Bayesian formulation of this model, prior distributions that represent *a priori* expectation, and associated uncertainty, are required for each model component. The model predictions (posterior distribution) are then obtained as the normalized product of the prior distribution and the likelihood, which is the contribution from fitting to the primary trial observations.
- Here, the prior distributions are obtained by fitting appropriate candidate survival models to relevant external data sources. Specifically, the Bayesian poly-hazard model is fitted according to the following procedure:

1. Estimating the prior distribution for the background hazard: a Gompertz distribution is fitted to trial-matched general population mortality data.[9]
2. Estimating the prior distribution for the disease-specific hazard: an appropriate hazard function (here, a log-normal distribution is used as an illustrative example) is estimated by fitting a poly-hazard model to relevant external data (here, we use registry data from the Surveillance, Epidemiology, and End Results (SEER) program[10]), using a fixed background hazard determined in step (1).
3. Estimating the posterior density: the model with prior distributions derived in steps (1) and (2) is fitted to the trial data using Monte Carlo sampling.[11]

- The performance of the Bayesian poly-hazard model fitted to data from CheckMate 227 is assessed by comparison to:
  - —Later observations from a more mature data cut of CheckMate 227 Part 1, with 61.3 months of minimum follow-up.[2]
  - —A standard parametric model (SPM) with post-hoc adjustment by general population mortality data.
  - —A Bayesian multi-parameter evidence synthesis[12] (B-MPES) model fitted using the same external data sources as the poly-hazard model, reported in a previous study.[13]

Results

- The prior and posterior survival functions for the overall population and disease-specific component in the Bayesian poly-hazard model differ dramatically (Fig. 1).
- Short-term extrapolations (e.g., at 5 years) from the Bayesian poly-hazard model are in agreement with later observations from the more mature trial data (Table 1).

Table 1. Milestone survival probability estimates (%) and 95% uncertainty intervals from parametric models fitted to the NIVO+IPI cohort in the primary data cut of CheckMate 227 Part 1, compared to trial observations from the later data cut (Kaplan-Meier estimate).

Time (years)	Survival probability estimates (%) by model			
	Bayesian poly-hazard	B-MPES	Adjusted SPM	Kaplan-Meier
5	20.7 [17.6-23.9]	22.4 [19.6-25.5]	21.0 [17.8-24.2]	22.5 [19.2-26.2]
10	9.1 [7.1-11.3]	10.7 [9.1-12.5]	11.1 [8.8-13.7]	-
15	4.3 [3.1-5.6]	4.4 [3.6-5.2]	7.4 [5.7-9.6]	-
20	1.8 [1.2-2.5]	1.4 [1.1-1.7]	4.4 [3.4-5.7]	-

- The Bayesian poly-hazard model predicts that the background and disease-specific hazards become equal at approximately 15 years, but the age-related mortality makes a non-negligible contribution to the overall hazard well before that point, increasing rapidly from around 8 years (Fig. 2).
- Long-term survival extrapolations from the Bayesian poly-hazard model are in close agreement with B-MPES predictions under a pessimistic scenario (from a previously reported model[13]) (Table 1, Fig. 3).
- The SPM (independent log-logistic distribution, selected naively based on statistical goodness-of-fit criteria) also yields agreement with the more mature trial dataset. However, on longer timescales, the SPM is more optimistic than both the Bayesian poly-hazard and B-MPES models.

Figure 1. Overall and component prior (above) and posterior (below) survival functions for the Bayesian poly-hazard model fitted to the NIVO+IPI cohort in the primary data cut of CheckMate 227 Part 1, compared to trial and SEER observations (Kaplan-Meier estimates).

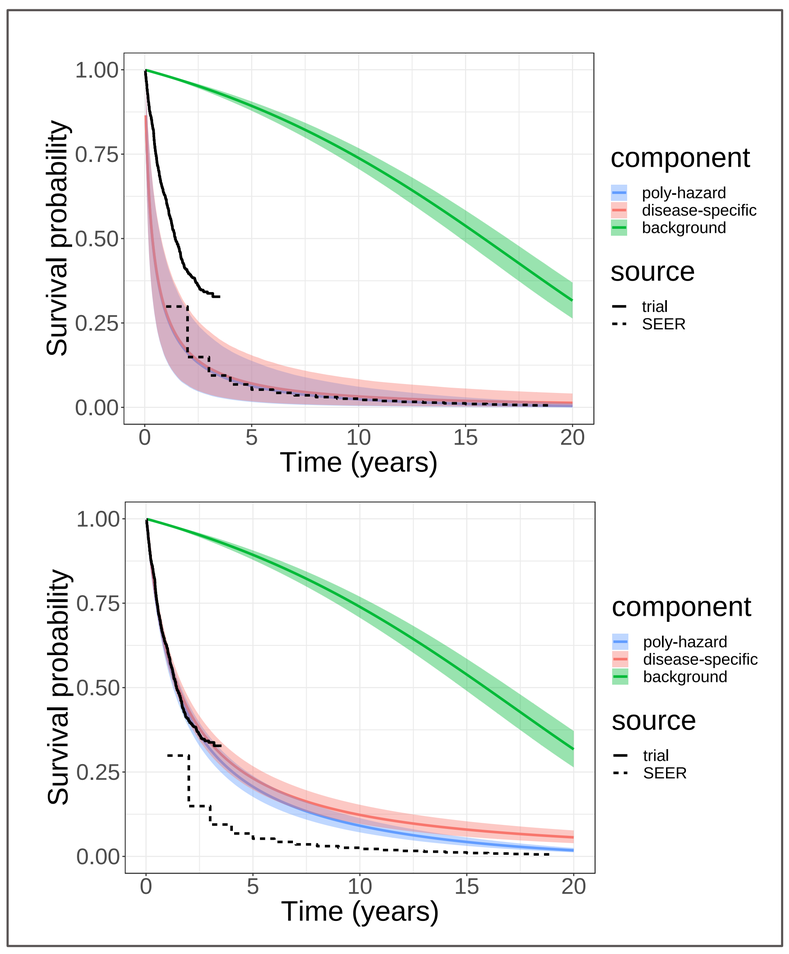


Figure 2. Overall and component posterior hazard functions for the Bayesian poly-hazard model fitted to the NIVO+IPI cohort in the primary data cut of CheckMate 227 Part 1, compared to trial observations (B-spline estimate).

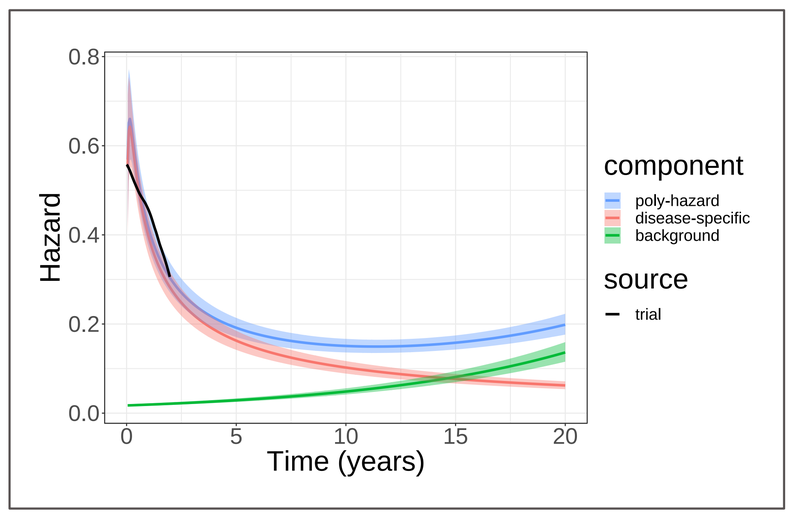
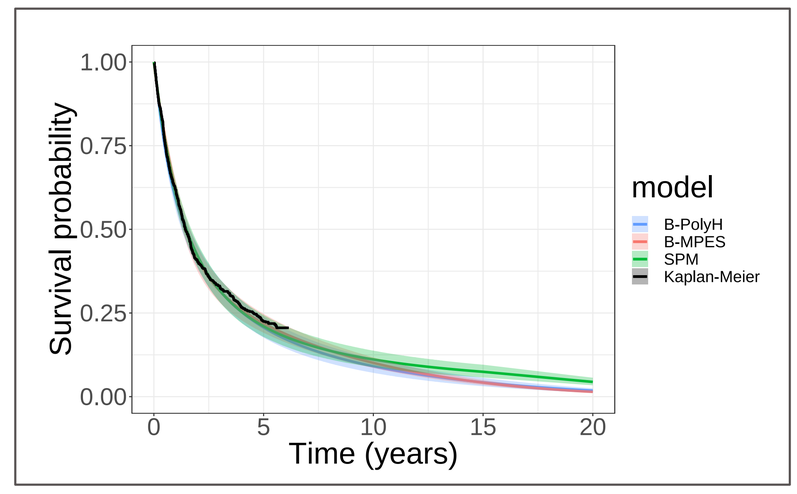


Figure 3. Survival functions of various parametric models fitted to the NIVO+IPI cohort in the primary data cut of CheckMate 227 Part 1, compared to trial observations (Kaplan-Meier estimate) from the later data cut.



- Adjusting the SPM by general population data from the point where these data are exceeded by predicted one-year conditional survival (15 years) tempers this effect after this timepoint, but a significant difference between the SPM predictions and estimates from the two Bayesian models still remains (Fig. 3).

Discussion

- The Bayesian poly-hazard model recovers a clinically plausible survival function in the posterior density, despite the use of SEER data, which is an overly pessimistic representation of *a priori* expectation for NIVO+IPI survival in CheckMate 227 Part 1, as it primarily reflects an outdated standard of care. These results suggest that the Bayesian poly-hazard approach is relatively robust to the choice of prior for the disease-specific hazard function. Further work will compare Bayesian poly-hazard models using different candidate distributions for the disease-specific hazard.
- The inclusion of the SEER and general population data ensures that the Bayesian poly-hazard model captures the qualitatively correct lifetime hazard trend and manifests a plausible, and not overly strong, survival plateauing effect beyond the trial follow-up period.
- The Bayesian poly-hazard model can be understood to be more conservative than the adjusted SPM not only through its use of SEER data, but also since it implicitly accounts for both general population and disease-specific mortality at all times. Thus, the poly-hazard model allows for residual disease-specific hazard after the point at which age-related mortality has become the larger contribution to the hazard.

Conclusion

- Bayesian poly-hazard models yield clinically plausible lifetime survival estimates for patients with aNSCLC receiving NIVO+IPI, even when follow-up is limited. Predictions are in close agreement with those from an alternative Bayesian model based on the same information sources (namely, a B-MPES model).
- Moreover, the Bayesian poly-hazard model has an appealing and clearly interpretable structure that aids validation of predictions. In addition, the direct integration of external data to support survival extrapolations makes the models more justifiable and reduces subjectivity, which is attractive in HTAs.
- Additional research is required to investigate this promising approach in other case studies, and to help define appropriate HTA guidelines for future use.
- Further work will investigate the use of alternative data sources to derive prior distributions for the disease-specific hazard, such as historical trial data or observational data that are specific to a novel therapy or class thereof; such data are likely to more closely reflect the treatment of interest and hence be less conservative than SEER.

References

1. Hellmann MD, et al. *N Engl J Med*. 2019; 381(21): 2020-2031.
2. Brahmer JR, et al. *J Clin Oncol*. 2023; 41(6): 1200-1212.
3. Chaudhary MA, et al. *Med Decis Making*. 2023; 43(1): 91-109.
4. Jackson C, et al. *Med Decis Making*. 2017; 37(4): 377-390.
5. Bullement A, et al. *Med Decis Making*. 2023; 43(5): 610-620.
6. Bartoš F, et al. *BMC Med Res Methodol*. 2022; 22(1): 238.
7. Rutherford MJ, et al. NICE DSU TSD 21. Published 2020.
8. Benaglia T, et al. *Stat Med*. 2015; 34(5): 796-811.
9. World Health Organization, WHO methods and data sources for life tables 1990-2019. WHO/DDI/DNA/GHE/2020.1. 2020.
10. Surveillance Research Program, National Cancer Institute. SEER*Stat. 8.4.0.1. 2022.
11. Carpenter B, et al. *J Stat Softw*. 2017; 76(1): 1-32.
12. Guyot P, et al. *Med Decis Making*. 2017; 37(4): 353-366.
13. Sharpe DJ, et al. Poster presentation at ISPOR 2023; May 2023; Boston, USA. Abstract MSR29.

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