Application of Bayesian multi-parameter evidence synthesis to achieve clinical plausibility of lifetime survival extrapolations in first-line advanced non-small cell lung cancer (aNSCLC)

MSR29

Daniel Joshua Sharpe,[1] Georgia Yates,[1] Llenalia-Maria García-Fernández,[2] Yong Yuan,[3] Adam L. Lee,[4] Mohammed Ashraf Chaudhary[3]

[1]Parexel, London, UK; [2]Parexel, Madrid, Spain; [3]Bristol-Myers Squibb, Lawrenceville, NJ, USA; [4]Bristol-Myers Squibb, Uxbridge, UK

Objectives

- To apply a flexible survival model that integrates relevant external data to improve the reliability of long-term survival predictions for patients with aNSCLC receiving nivolumab plus ipilimumab, using the Bayesian multi-parameter evidence synthesis (B-MPES) approach. The estimates are compared to those from uninformed standard parametric models (SPMs).

Background

- CheckMate 227 Part 1 is a phase 3 randomized trial in patients with stage IV or recurrent NSCLC, in the first-line setting[1]. Here, we combined the pair of arms for patients randomized to nivolumab plus ipilimumab (NIVO+IPI), and likewise pooled the pair of chemotherapy (CHEMO) arms, to estimate time-to-event models for overall survival (OS) by treatment, irrespective of tumor PD-L1 expression level. The dataset comprises 1166 patients in total
- Durable benefit has been demonstrated with NIVO+IPI, hence, there is the potential for long-term survivorship from this regimen[2]
- SPMs are typically unable to capture complex survival effects such as heterogeneous response and age-related mortality, which may manifest as multimodal hazard functions. Thus, where there is appreciable survival on longer timescales, more flexible parametric models that can accommodate these effects are desirable, and naïve extrapolation is inadvisable[3]
- B-MPES has been proposed as a flexible method that can improve reliability of long-term survival estimates by integrating external data sources to inform extrapolations beyond the trial follow-up period[4]
- In previous work, we have shown that B-MPES models fitted to data from trials of second-line nivolumab in aNSCLC demonstrate clinically reasonable and stable estimates for long-term OS[5]

Methods

- B-MPES models based on underlying independent two-knot natural cubic spline functions smoothing the cumulative hazard functions, and a set of seven candidate SPMs recommended by NICE, were fitted to the 2- (primary) and 5-year database locks (DBLs) of CheckMate 227 Part 1, with 29.3[1] and 61.3 months[2] of minimum follow-up, respectively.
- The B-MPES model simultaneously fits the NIVO+IPI and CHEMO arms. External data directly influence the CHEMO survival function, and indirectly influence the NIVO+IPI survival function via a hazard ratio condition
- The B-MPES model uses vague normal priors. External data instead inform the model via additional contributions to the posterior density
- Additional binomial contributions to the posterior density are applied to the CHEMO arm based on one-year conditional survival derived from relevant registry data extracted from the Surveillance, Epidemiology, and End Results (SEER) program[7] and from trial-matched World Health Organization general population mortality data[8]
- The registry data comprises all available observations from the SEER database[7] for patients with stage IIIB or IV lung cancer exhibiting either squamous or non-squamous histology and has 20 years of follow-up. The registry data are applied for the time period 3 to 20 years (for the 2-year DBL) or 6 to 20 years (for the 5-year DBL)
- The one-year conditional survival probabilities derived from general population data are applied for the time period 28 to 30 years. These data are included to ensure that the model appropriately captures an age-related mortality effect
- The external data indirectly informs the NIVO+IPI arm via a contribution to the posterior density from a normal distribution representing a hazard ratio condition. Here, we considered a conservative scenario where the hazard ratio is conditioned to be equal to one from 11 years onwards.
- Internal knots in the underlying spline functions of B-MPES are located at half the minimum follow-up time and at half the maximum model time (i.e., 15 years)
- The B-MPES model was fitted by the No U-Turn Sampler (NUTS) procedure using the Stan program.[9] Adequate convergence was confirmed by calculation of the potential scale reduction factors and effective sample sizes. B-MPES estimates are reported as posterior means with uncertainty quantified by 95% credible intervals (CrIs)
- SPMs were fitted by maximum-likelihood estimation. The best-fitting SPMs were chosen naively as those with the lowest Akaike information criterion (AIC) metric and were neither validated nor adjusted using external data. Uncertainty was quantified by 95% confidence intervals (CIs)
- Survival estimates in the NIVO+IPI arm were assessed by survival probability at 5-, 10-, 15-, and 20-years and restricted mean survival times (RMSTs) at 5- and 20-years

Table 1. Milestone survival probabilities (%) in the NIVO+IPI arm, and associated uncertainty, for models fitted to the 2- and 5-year database locks (DBLs) of CheckMate 227 Part 1.

Time (years)	2-year DBL		5-year DBL		
	B-MPES	SPM	B-MPES	SPM	KM
5	22.4 [19.6-25.5]	21.0 [17.8-24.2]	21.4 [18.6-24.4]	21.6 [18.7-24.5]	22.5 [19.2-26.2]
10	10.7 [9.1-12.5]	11.1 [8.8-13.6]	10.3 [8.8-12.0]	11.6 [9.5-13.8]	-
15	4.4 [3.6-5.2]	7.4 [5.6-9.5]	4.3 [3.6-5.1]	7.8 [6.2-9.6]	-
20	1.4 [1.1-1.7]	5.5 [4.1-7.2]	1.4 [1.1-1.6]	4.7 [3.6-6.0]	-

KM denotes Kaplan-Meier estimator. Model estimates from B-MPES are posterior means and uncertainty is quantified by 95% CrIs. Uncertainty in estimates from SPMs is quantified by 95% CIs.

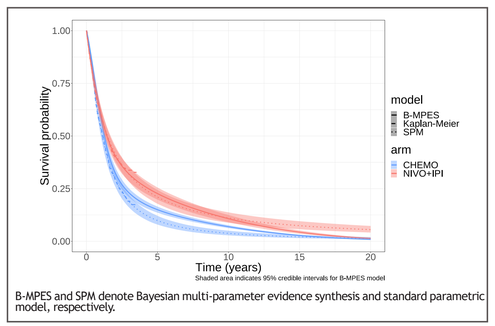
Results

- The best-fitting (i.e., lowest-AIC) SPMs for the 2- and 5-year DBLs are both based on independent log-logistic distributions
- The B-MPES model and SPM estimated from the 2-year DBL are in close agreement within the follow-up period and for short-timescale extrapolations up to around 10 years. At longer times, the predictions begin to diverge, with the B-MPES model yielding a significantly more conservative estimate [Fig 1]
- The B-MPES model fitted to the 2-year DBL accurately anticipates the observed NIVO+IPI survival in the extended follow-up of the 5-year DBL (5-year OS: 22.4% [95% CrI: 19.6-25.5%]) vs 22.5% [95% CI: 19.2-26.2%] Kaplan-Meier) [Tables 1 & 2]
- The best-fitting SPM estimated from the 2-year DBL similarly yields agreement with subsequently observed 5-year OS (21.0% [95% CI: 17.8-24.2%])
- Long-term extrapolations from B-MPES are conservative and are consistent between the earlier and later DBLs. This is best seen from the predictions for 20-year OS: 1.4% [95% CrI: 1.1-1.7%] 2-year DBL vs 1.4% [95% CrI: 1.1-1.6%] 5-year DBL).
- In contrast, long-term estimates from SPMs are more sensitive to successive follow-up, are more optimistic, and have higher uncertainty (e.g., 20-year OS: 5.5% [95% CI: 4.1-7.2%] 2-year DBL vs 4.7% [95% CI: 3.6-6.0%] 5-y DBL)

Discussion

- The B-MPES model employed herein is sufficiently flexible to accommodate the multiple external data sources, and thus yields a more clinically plausible model that correctly captures the trend in conditional survival probabilities on a lifetime scale [Fig 2]

Figure 1. Observed and estimated survival functions for the 2-year database lock of CheckMate 227 Part 1.



B-MPES and SPM denote Bayesian multi-parameter evidence synthesis and standard parametric model, respectively.

- The B-MPES model therefore appropriately represents survival effects that dominate the observed trend at longer times, such as age-related mortality
- The stability and conservativeness of B-MPES estimates are a direct result of incorporating external information, which ensures that the survival extrapolations are not naïve and that the conditional survival probabilities decrease at longer times in accordance with the real-world data
- Conversely, SPMs impose overly restrictive shapes on the hazard function (e.g., here, the SPM hazard functions are strictly unimodal). This lack of flexibility enforces conditional survival probabilities that monotonically increase after the follow-up period and eventually exceed general population survival
- Thus, SPMs typically require some form of post-hoc adjustment and external validation in the selection criteria in order to yield clinically plausible estimates for long-term survival. Short-term extrapolations may be more reasonable, but it is nonetheless desirable to validate such predictions with available external data
- The restrictive assumptions imposed by the SPMs are further evident in the NIVO+IPI vs CHEMO hazard ratios, which are close to time-independent in the SPM case but are strongly time-varying in B-MPES [Fig 3]
- The B-MPES model described herein is highly conservative for NIVO+IPI survival since the hazard ratio condition imposes a somewhat strict treatment waning scenario and the registry data reflects less effective and outdated treatment regimens. Scenario analyses could explore less pessimistic alternative schemes
- The implicit integration of external data sources into the B-MPES model is desirable in the context of health technology assessments, where models should be transparent and their projections justifiable
- Further work will examine the incorporation of relevant supplementary trial data as a further source of information that can be leveraged within the B-MPES framework, and assess the impact of model specification through sensitivity and scenario analyses surrounding choice of parametric function, the timepoints at which external data are applied, and the hazard ratio condition

Table 2. Restricted mean survival time (months) in the NIVO+IPI arm, and associated uncertainty, for models fitted to the 2- and 5-year database locks (DBLs) of CheckMate 227 Part 1.

Time (years)	2-year DBL		5-year DBL		
	B-MPES	SPM	B-MPES	SPM	KM
5	26.1 [24.3-27.9]	25.9 [24.1-27.8]	26.2 [24.4-28.0]	26.2 [24.4-27.9]	26.2 [24.4-28.1]
20	41.6 [37.7-45.7]	44.2 [38.8-49.8]	41.1 [37.3-45.0]	45.3 [40.4-50.6]	-

KM denotes Kaplan-Meier estimator. Model estimates from B-MPES are posterior means and uncertainty is quantified by 95% CrIs. Uncertainty in estimates from SPMs is quantified by 95% CIs.

Figure 2. One-year conditional survival probabilities estimated from parametric models fitted to the 2-year database lock of CheckMate 227 Part 1, compared to Kaplan-Meier estimates and estimates from external data.

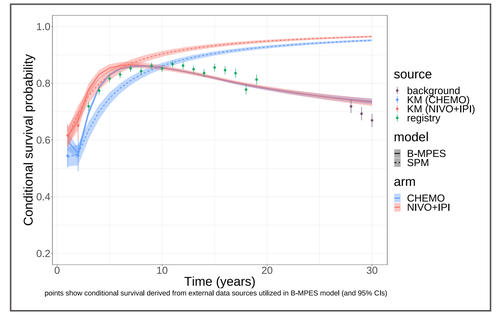
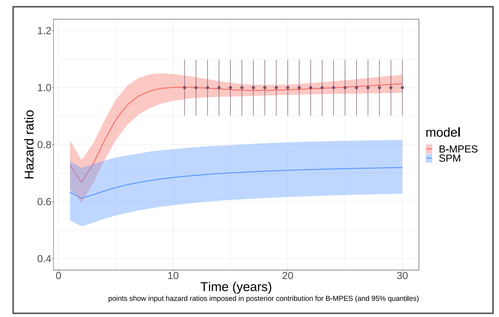


Figure 3. Predicted hazard ratios from parametric models fitted to the 2-year database lock of CheckMate 227 Part 1. Distributions for the expected hazard ratio at selected times that are used as inputs in B-MPES are also shown.



Conclusions

- With appropriately selected external data and model assumptions, B-MPES models can generate robust and clinically plausible estimates of long-term OS for patients with aNSCLC receiving NIVO+IPI, even when follow-up is limited.

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. Bullement A, Latimer NR, and Bell Gorrod H. *Value Health* 2019; 22(3): 276-283.
4. Guyot P, et al. *Med Decis Making* 2017; 37(4): 353-366.
5. Chaudhary MA, et al. *Med Decis Making* 2023; 43(1): 91-109.
6. Latimer N. NICE DSU Technical Support Document 14. Published 2013.
7. Surveillance Research Program, National Cancer Institute. SEER*Stat. 8.4.0.1. 2022.
8. World Health Organization, WHO methods and data sources for life tables 1990-2019. WHO/DDI/DNA/GHE/2020.1. 2020.
9. Carpenter B, et al. Stan: A Probabilistic Programming Language. *J Stat Softw* 2017; 76(1): 1-32.

Presented at ISPOR 2023; May 7-10, 2023; Boston, MA

Email: Ashraf.Chaudhary@bms.com

Copies of this poster are for personal use only and may not be reproduced without written permission of the authors.