PCN422

BAYESIAN MODEL AVERAGING FOR EXTRAPOLATING SURVIVAL: A PAN-TUMOR PERSPECTIVE ON NIVOLUMAB AND IPILIMUMAB

Sven L. Klijn[1], Katharina Nickel[2], Elisabeth Fenwick[3], Rachel Bakker[1], Bill Malcolm[4]

[1]Pharmerit International, Rotterdam, The Netherlands; [2]Pharmerit International, Berlin, Germany; [3]Pharmerit International, Oxford, United Kingdom; [4]Bristol-Myers Squibb Pharmaceuticals Ltd, Uxbridge, United Kingdom

Introduction

Background

- Survival in oncology is commonly extrapolated using traditional parametric distributions or cubic splines.[1]
- However, there is a challenge in determining the most appropriate extrapolation form.
  - Guidance on selection methods of extrapolations is not always conclusive.
  - This has become an issue especially for Immuno-Oncology (IO) therapies or other therapies which might show different survival patterns compared to chemotherapies, i.e. exhibiting a long tail in Overall Survival (OS).
- Bayesian model averaging (BMA) can provide a comparative assessment of goodness-of-fit (GOF) of individual extrapolations that is easily interpretable and can be combined into a single model.
  - Further, BMA does not force the modeler to select one distribution per se but rather to empirically acknowledge the underlying heterogeneity in the data by allowing all relevant distributions to contribute to an aggregated form concomitantly.
  - Selection of the extrapolation form is frequently scrutinized in HTA procedures. Leveraging BMA allows the evaluation of a composite estimate as well as meaningful quantification of the GOF of individual forms.

Objective

- The objective of this study was to explore patterns in GOF of traditional parametric extrapolations and cubic splines with BMA, applied to data for nivolumab (NIVO) and ipilimumab (IPI) across four indications.

Methods

Data Input

- Indication and treatment arms assessed are listed in Table 1.

Table 1. Data Sources per Indication

Indication	Treatment arms	Outcome of interest	Key clinical trial
Adjuvant melanoma	Nivolumab, Ipilimumab	Relapse free survival (RFS)	CheckMate 238 trial; 24-months follow-up data[2]
Squamous cell cancer of the head and neck (SCCHN)	Nivolumab	Overall survival	CheckMate 141 trial; 24-months follow-up data[3]
Metastatic melanoma	Nivolumab, Nivolumab + Ipilimumab, Ipilimumab	Overall survival	CheckMate 067 trial; 4-years follow-up data[4]
Squamous non-small cell lung carcinoma (NSCLC)	Nivolumab	Overall survival	CheckMate 017 trial; 4-years follow-up data[5]

Data Preparation

- Kaplan-Meier (KM) curves were digitized using *WebPlotDigitizer*, a validated open-source software package.[6,7] Accuracy of digitized curves were compared to published landmark survival probabilities.
- After digitization of the KM curves, patient level data was simulated using the algorithm developed by Guyot et al. (2012).[8] Accuracy of simulated survival was compared to published landmark survival probabilities.
- Thirteen independent parametric extrapolations - 7 standard distributions (Exponential, Gamma, Generalized gamma, Gompertz, Loglogistic, Lognormal, Weibull) and 6 splines (1-knot hazard spline, 1-knot probit spline, 1-knot odds spline, 2-knots hazard spline, 2-knots probit spline, 2-knots odds spline), were fitted to the curves estimating survival beyond the trial periods.
- The curves were fitted using the R package *flexsurv* and the GOF of each distribution was estimated via the Bayesian Information Criterion (BIC).
- The extrapolated survival curves were not a djusted for general population mortality.
- No assumptions were made regarding proportional hazards, as survival of each treatment arm was extrapolated independently.

Bayesian Model Averaging

- BMA involves building a weighted mean of the quantity of interest (here survival) from parametric models, where the weights are determined as a function of the BIC associated with each model and a prior probability that each model is the true model.
- Following the methodology from Negrín et al. (2017)[9] the posterior probability p(S=s | D) of each distribution being true was estimated via:

$$p(S = s|D) = \frac{\exp(-0.5BIC_s)\, p(S = s)}{\sum_{s=1}^{S} \exp(-0.5BIC_s) p(S = s)}$$

- As such the posterior distribution depended only on the BIC of each distribution (BIC_s) and its prior p(S=s).
- Prior probabilities for each survival model s=1,2,…,S, were set to the uninformative prior 1/S.
- Posterior probabilities of the 13 parametric extrapolations were compared across indication, treatment, and outcome measurement.

Results

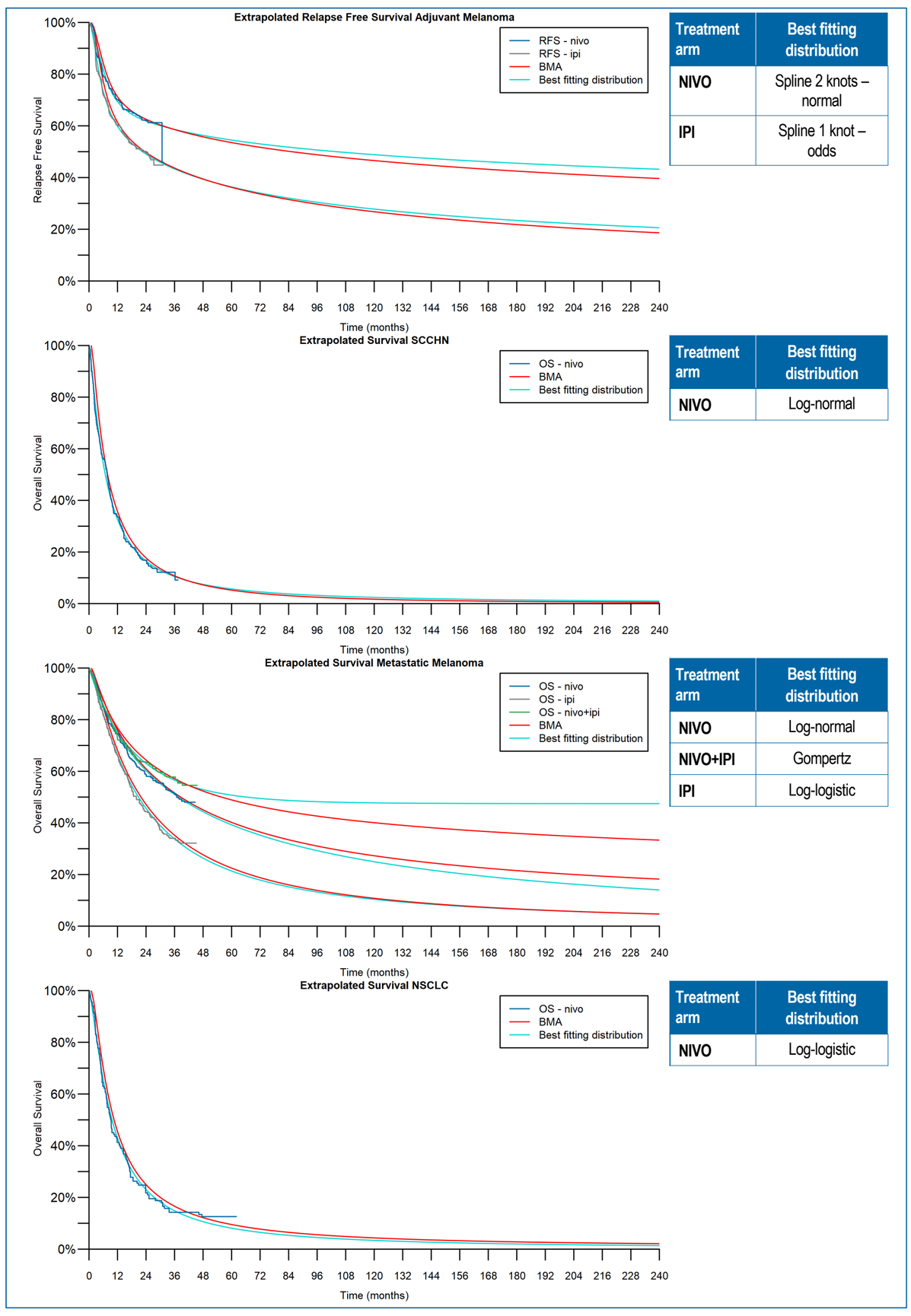
Data Preparation

- The digitization was found to be precise with an absolute difference between digitized curve and published values <1.5%.
- Visual inspection of the simulated patient level data demonstrated a good fit to the published curves and the comparison to published survival probabilities indicated only minor (<1.5%) deviations of the simulated survival from the published data.

Bayesian Model Averaging

- Posterior probabilities of the 13 parametric extrapolations per indication and treatment are presented in Table 2. The observed survival, BMA and best fitting distribution per indication are presented in Figure 1. The large step in the RFS KM curve is due to low numbers at risk towards maximum follow-up.
- Generalized gamma, Gompertz, Log-logistic and Log-normal distributions mostly induced long-term survival tails and had higher posterior probabilities.
- The Log-normal distribution was among the best-fitting distributions for OS data across all indications and arms, with posterior probabilities ranging from 18%-70%.
- Splines provided a good fit to RFS with >98% of the posterior probability attributed to a spline.
- For OS, 1-knot splines provided better fits than 2-knot splines, although cumulative posterior probabilities of splines were <31%.

Figure 1. Plotted observed Survival, BMA, and Best Fitting Distribution



Discussion

- Comparison across indications and treatment arms demonstrated that the Exponential, Gamma, and Weibull distributions with a survival trend towards zero, were a bad fit (posterior probabilities <1%) for extrapolating survival data of patients treated with NIVO and/or IPI.
- Generalized gamma, Gompertz, Log-logistic and Log-normal distributions all provided better fit and mostly induced long-term survival tails consistent with the mechanism of action of the IO therapeutic class where the aim is longer term survival and cure.
- The GOF of distributions differed between the outcomes of RFS and OS:
  - For the OS data the Log-normal distribution was among the best-fitting distributions for OS data across all indications and arms.
  - The RFS data was best described by splines.
- More mature OS trial data seemed to be described best by a mixture of Log-logistic and Log-normal distributions while less mature OS data was described best by Log-normal and Gompertz distributions.
- Differences in best fitting functional forms between RFS and OS require further investigation.

Strengths

- This analysis was based on a retrospective assessment of clinical trials with a long follow up (>2 years).
- Trial data with different maturity levels have been evaluated (between <10% and >50% survival).
- This assessment produced insights into the general GOF of distributions to survival data of patients treated with the immuno-oncological treatments NIVO and/or IPI across a range of indications.

Limitations

- The GOF of each distribution was only based on statistical fit not including clinical information.
- The results of the analyses could not be validated using external data.
- This study examined four different indications, however, generalizability of the results beyond these indications has not been evaluated and may be limited. This holds especially true for RFS for which only one indication has been studied.
- The analyses were based on digitized and simulated data instead of patient-level data, however, accuracy of the digitization and simulation steps has been assessed and differences were minor.

Conclusions

- The BMA approach has been used to evaluate the GOF of individual distributions to survival data of patients treated with the immuno-oncological treatments NIVO and/or IPI across a range of indications.
- Exponential, Gamma and Weibull distributions provided a poor GOF to NIVO and IPI survival across indications.
- For OS, Generalized gamma, Gompertz, Log-logistic and Log-normal distributions provided a good fit to observed data and may thus be a preferable choice for survival extrapolations that naturally induce a plateau.
- The analyzed RFS data was best described by splines, indicating potential differences in the extrapolations of OS and RFS efficacy outcomes.

References

1. Woods B, et al. NICE DSU Technical Support Document 19. Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review. 2017.
2. Weber J, et al. Adjuvant therapy with nivolumab (NIVO) versus ipilimumab (IPI) after complete resection of stage III/IV melanoma: Updated results from a phase III trial (CheckMate 238). *J Clin Oncol* 2018; 36:9502-9502.
3. Ferris R, et al. Nivolumab vs investigator's choice in recurrent or metastatic squamous cell carcinoma of the head and neck: 2-year long-term survival update of CheckMate 141 with analyses by tumor PD-L1 expression. *Oral Oncol* 2018;81:45-51.
4. Wolchok J, et al. Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. *N Engl J Med* 2017;377: 1345-1356.
5. Scott J, et al. Four-year survival with nivolumab in patients with previously treated advanced non-small-cell lung cancer: a pooled analysis. *Lancet* 2019;201395-1408.
6. Drevon D, et al. Intercoder Reliability and Validity of WebPlotDigitizer in Extracting Graphed Data. *Behav Modif* 2017;41:323-339.
7. Moeyaert M, et al. Reliability, Validity, and Usability of Data Extraction Programs for Single-Case Research Designs. *Behav Modif* 2016;40:874-900.
8. Guyot P, et al. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;12:9.
9. Negrin M, et al. Bayesian Solutions for Handling Uncertainty in Survival Extrapolation. *Med Decis Making* 2017;37:367-376.

Acknowledgments

- Bristol-Myers Squibb (Princeton, NJ) and ONO Pharmaceutical Company Ltd. (Osaka, Japan)
- This study was supported by Bristol-Myers Squibb
- All authors contributed to and approved the poster.

Table 2. Bayesian Model Averaging Results

Distribution	Posterior probabilities of the distribution [%]						
	RFS data		OS data				
	Adjuvant melanoma - nivolumab	Adjuvant melanoma - ipilimumab	SCCHN - nivolumab	Metastatic melanoma - nivolumab	Metastatic melanoma – nivolumab + ipilimumab	Metastatic melanoma - ipilimumab	NSCLC - nivolumab
Exponential	<0.01	<0.01	<0.01	0.18	<0.01	0.62	<0.01
Gamma	<0.01	<0.01	<0.01	0.08	<0.01	0.04	<0.01
Generalized gamma	<0.01	0.07	4.53	2.80	5.00	2.34	2.74
Gompertz	1.84	<0.01	1.49	**20.58**	**41.02**	2.58	7.48
Log-logistic	<0.01	<0.01	**15.69**	8.21	1.20	**44.67**	**40.75**
Log-normal	<0.01	<0.01	**69.93**	**47.82**	**28.48**	**36.36**	**18.37**
Weibull	<0.01	<0.01	<0.01	0.18	0.04	0.05	<0.01
Spline 1 knot – hazard	**24.65**	**43.00**	1.42	7.44	**8.43**	2.56	7.97
Spline 1 knot – normal	<0.01	1.62	**4.62**	2.97	5.32	2.48	2.50
Spline 1 knot – odds	**22.45**	**50.49**	1.2	**8.31**	8.34	**7.10**	**15.04**
Spline 2 knots – hazard	5.91	2.16	0.16	0.49	0.68	0.32	1.57
Spline 2 knots – normal	**33.88**	0.46	0.39	0.42	0.83	0.41	1.68
Spline 2 knots – odds	11.26	**2.20**	0.54	0.52	0.67	0.48	1.90

Note: The posterior probabilities of the three distributions with the highest weight per indication and treatment arm are printed bold.
NSCLC: Squamous Non-small Cell Lung Cancer ; OS: Overall Survival ; RFS: Recurrence-free Survival; SCCHN: Squamous Cell Carcinoma of the Head and Neck.

Presented at ISPOR Europe 2019; November 2–6, 2019; Copenhagen, Denmark

Email: sklijn@pharmerit.com