HTA72

Survival Extrapolation for Immuno-Oncology Drugs: Pessimism or Optimism? a Targeted Literature Review and a Validation-Based Case Study

Phung TH[1]

[1] Centre for Health Economics, University of York, United Kingdom

Acknowledgments: I gratefully acknowledge the guidance and support of Matthew Taylor, Karin Butler, Erin Barker, Joe Moss at York Health Economics Consortium and Pedro Saramago at the University of York for this research

Background

The objectives of this study are to (1) examine the justification of selecting complex survival extrapolation models among ones suggested in the DSU Technical Support Document 21[1] and how Evidence Review Groups and appraisal committees responded to the use of complex models in the UK, (2) assess the accuracy of the used and unused methods in the NICE technical appraisal 581[2] (ipilimumab plus nivolumab in renal cell carcinoma) to establish whether alternative methods could have provided better prediction.

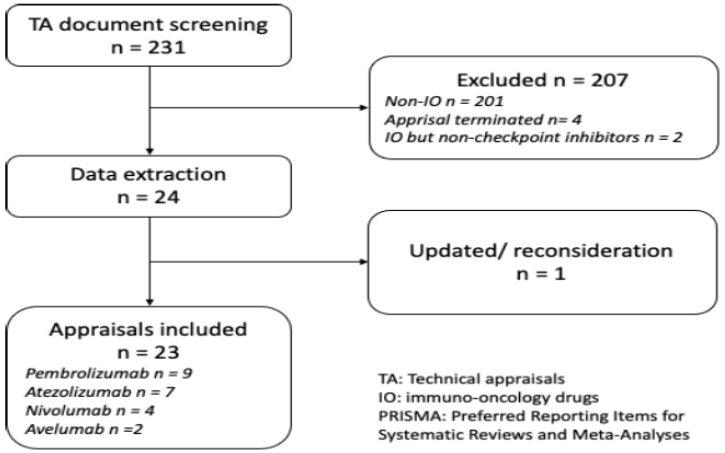
Methods

All NICE TAs for immune-oncology drugs (IOs) conducted from January 2018 to July 2021 were included. Information of justifications for extrapolation methods, critiques of ERGs, and final appraisal determination of committees over these methods were extracted.

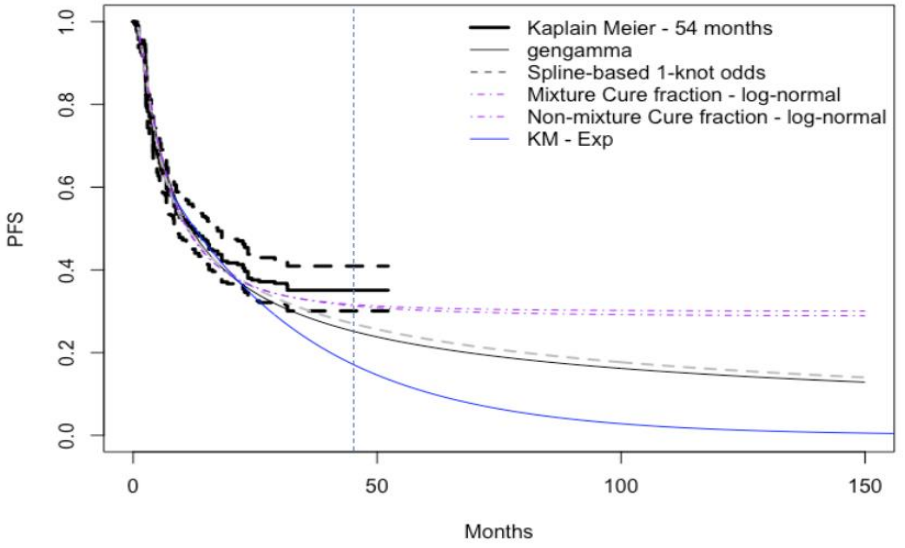
A range of extrapolation methods including fully-fitted parametric, semi-parametric, cubic spline-based, and cure fraction models were performed using the 2-year dataset of the pivotal trial in TA581 (CHECKMATE 214[3]). Best statistical fit models for each method were validated by the 4-year dataset[4] using both landmark survivals and restricted mean survival time.

Results

A total of 23 TAs were included. Once the use of standard-parametric models was considered inappropriate (n=16), the majority of TAs (n=14/16) failed to appropriately justify selecting a complex model among alternatives. ERGs and committees generally agreed on the methods modelling observed hazards (semi-parametric, cubic spline-based models) whilst against modelling the unobserved hazards (response-based and cure fraction models). The response-based model was employed in one TA, but neither the ERG nor the NICE committee accepted its use. All TAs that incorporated a "cured" proportion in their models (n=3) was rejected.



In the validation-based case study, only cure fraction models provided accurate 4-year survival predictions (for both OS and PFS) compared to the updated dataset. All other alternatives including models used in TA581 by the company and ERG underestimated 4-year survivals.



Conclusion

Though the targeted literature review showed conservative standpoints from ERGs and NICE committees over response-based and cure fraction models, the validation-based case showed that cure fraction models could provide more accurate long-term survival predictions. This encourages future TAs to collect further long-term information to reassess previous model findings.

References: 1. Rutherford et al. 2020. "NICE DSU Technical Support Document 21. Flexible Methods for Survival Analysis"; 2. National Institute for Health and Care Excellence 2019. Nivolumab with ipilimumab for untreated advanced renal cell carcinoma. Technology appraisal guidance [TA581]; 3. Motzer et al. 2018. "Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma." The NEJM 378 (14): 1277–90; 4. Albiges et al. 2020. "Nivolumab plus Ipilimumab versus Sunitinib for First-Line Treatment of Advanced Renal Cell Carcinoma: Extended 4-Year Follow-up of the Phase III CheckMate 214 Trial." ESMO Open 5 (6): e001079