MSR141

Identifying and extrapolating immature OS data in survival analysis: A review of the variation in methods used in oncology STAs for NICE

Moura, A[1], Westerberg, E[1], Cheah, Z[2], Ho, M[2], Simons, C[2]

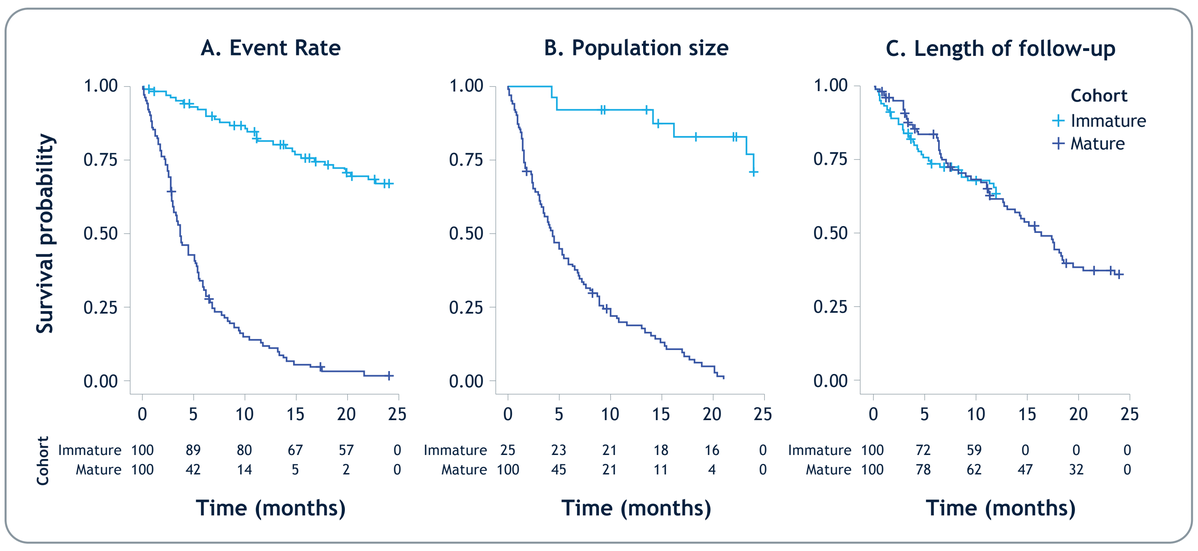
[1]OPEN Health HEOR & Market Access, Rotterdam, The Netherlands
[2]OPEN Health HEOR & Market Access, London, UK

OPEN HEALTH

INTRODUCTION

- Technology appraisals (TAs) submitted to the National Institute for Health and Care Excellence (NICE) describe cost-effectiveness analyses conducted on medical therapies, including survival analyses carried out on time-to-event data.
- Due to time and cost constraints surrounding clinical trials, survival analyses conducted in NICE TAs often involve working with immature overall survival (OS) data, where few events are observed over the course of patient survival within the trial duration.
- There is no apparent clear definition in the literature of what defines immature data in the context of survival, time to event data. However, a common definition is often that the observed median survival time has not been reached.
- There are several factors that can influence the maturity of OS data, including duration of follow-up, sample size, and event rates as shown below in Figure 1.
- Extrapolations of survival data are often carried out to obtain long-term estimates of effectiveness and survival for incorporation into cost-effectiveness models. However, in the context of immature OS data, additional care should be taken in interpretation of these extrapolated curves due to the increase in uncertainty surrounding these long-term values.
- Despite the presence of NICE DSU TSD 14[1] and 21[2] which outline the approaches for extrapolating incomplete survival data, there is currently no explicit guidance from NICE on how to deal with immature OS data (where the observed KM has not reached its median). As a result, the extrapolation methods for immature OS data can vary widely between submitted TAs.

Figure 1. Mature and immature OS, simulated example data. Differences in maturity can be attributed to alterations in event rate (A), population size (B), and length of follow-up (C)



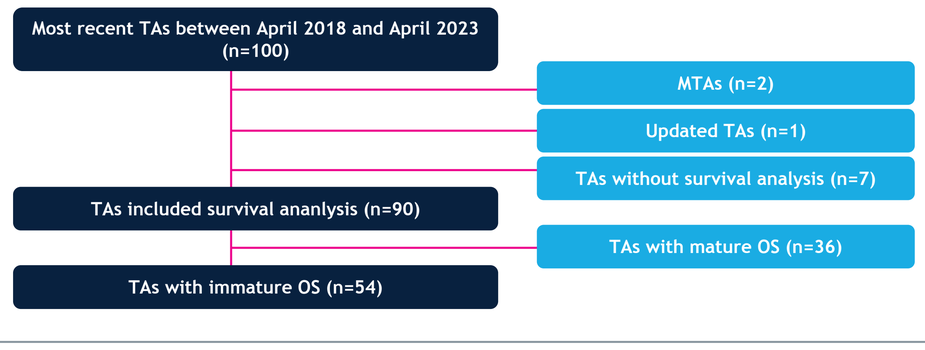
OBJECTVES

- The aim of this work was to explore the variety of statistical methods used to extrapolate immature OS data and how these methods were viewed by the NICE external assessment group (EAG) or evidence research group (ERG) committees, with the aim of determining whether guidelines on extrapolation methods for immature OS data should be more firmly outlined by NICE documentation.

METHODS

- All publicly available NICE single technology assessments (STAs) in oncology that were published on the NICE website from April 2018 up to April 13, 2023, were identified.
  - If an STA contained the word ‘cancer’, ‘carcinoma’, ‘myeloma’, ‘chemotherapy’, ‘lymphoma’, ‘leukaemia’, or ‘melanoma’, it was flagged as an oncology STA in Excel.
  - STAs that did not contain these words were recorded as ‘Not available’ (N/A) and were checked to ensure that no oncology STAs had been missed by applying the algorithm above.
- When OS data in the submission was described as mature by the company or no mention of immature OS data was found within the submission document, information on the approach used to extrapolate OS data was not extracted.
- Information on the maturity of OS data, approach used to extrapolate OS data, and EAG/ERG opinion on the related methods used were extracted and analysed in a descriptive manner.
- All data were extracted and analysed descriptively by one reviewer, with a sample reviewed by a second reviewer.
  - All authors listed contributed to either the extraction or reviewing of data.
- Requests for unavailable supporting documents were not made.

Figure 2: Flow chart showing the identification of NICE TAs including immature OS data



Abbreviations: MTA, multiple technology appraisal; OS, overal survival; TA, technology appraisal.

RESULTS

- Out of all the STAs identified, the 100 most recent STAs were screened (2020–2023) and selected based on the criteria outlined in Figure 2, where multiple technology assessments (MTAs) and a TA (TA519) that was replaced with a TA with updated guidance was removed. Following this, TAs that did not extrapolate OS data with survival analysis were excluded.
- Out of 90 STAs that included survival analysis with extrapolation of OS data, 60.0% (n=54) stated that they had immature data with a total median follow-up duration of 20.6 months (interquartile range [IQR]: 16.4–28.1). Conversely, 40.0% (n=36) of STAs stated that they had mature data, with a total median follow-up duration of 20.5 months (IQR: 14.9–45.5).
- Of those who stated that they were using immature data, 77.8% (n=42) of studies did not reach median OS. Of the 22.2% (n=12) studies that stated immature data but met median overall survival, information on either the median follow-up or median OS was redacted in 9 studies.
- The most commonly employed methods used to extrapolate immature OS data were standard parametric models (87.5%, n=46) and spline models (19.6%, n=11). As submissions were able to fit survival curves based on multiple different methods the values in Figure 3 sum to more than 100%.
- It is of note that of the 5 STAs that fitted mixture-cure models, 2 of them received the critique that the EAG/ERG did not consider the cure assumption to hold on the basis of the available data.
- The most common rationales behind the EAG/ERG agreeing or disagreeing with the base-case approach to extrapolation can be found in Figure 4.

Figure 3. Proportion of methods used to extrapolate immature OS data outlined in submissions

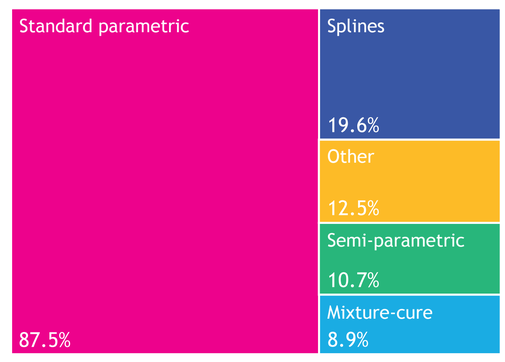


Figure 4. Reasons for why the EAG/ERG agreed or disagreed with methods used to extrapolate immature OS data

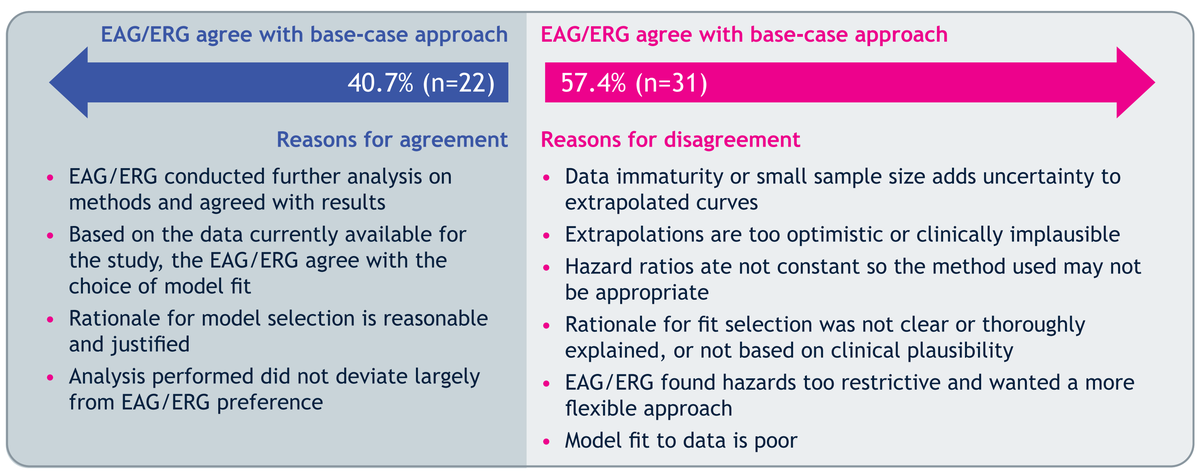
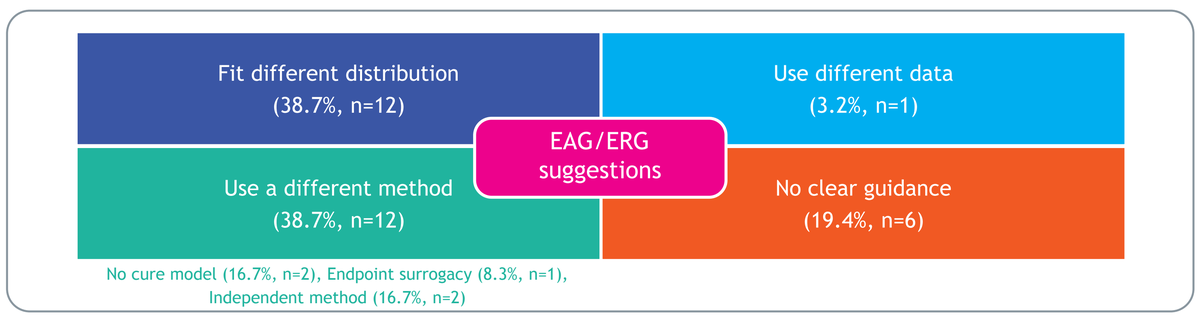


Figure 5. Advice given by the EAG/ERG on how best to extrapolate immature OS data when they disagreed with the methods employed in the original submission



CONCLUSIONS

- Based on our findings, we found that 60.0% of TAs submitted to NICE claimed to have immature OS data.
- 85.2% of TAs used standard parametric models to extrapolate immature OS in the survival analyses conducted, with 59.8% of TAs only fitting parametric survival curves.
- Out of all the TAs evaluated, the EAG/ERG did not agree with the selection of the base case survival extrapolation used in 57.4% of them.
- When the EAG/ERG disagreed with the base case model choice used to extrapolate immature OS data, 77.4% of advice given was to alter the choice of model or parametric distribution fit used in extrapolations in line with NICE DSU TSD 14[1] and 21.[2]
- In 19.4% of submissions in which the EAG/ERG disagreed with the base case models chosen for the approach used to extrapolate immature OS data, no explicit advice on how to adjust the methods or base case decision was outlined by the EAG/ERG.
- In 40.0% (n=2 out of 5) of the STAs that fitted mixture-cure models for extrapolation, the EAG/ERG critiqued the use of such models on immature data.
- There are several limitations of this study, many of which motivate additional work. This review focussed only on analysing the approaches of extrapolating OS in those submissions which claimed to have immature data. It would be of interest to see how these critiques by the EAG and methodologies also apply to those submissions which did not make such a claim on data immaturity and/or on other endpoints such as progression free survival.
- Our results show that over half of the NICE oncology submissions that were reviewed included extrapolation of immature OS data (which either did not reach median survival in the observed time period and/or claimed that their data was immature), and that the EAG/ERG disagreed with the choice of base case statistical distribution and/or methodology used in over half of these submissions.
- A frequent reason for disagreement was around the plausibility of the long-term extrapolations and their relationship with the clinical experts’ expectations. This suggests that new methods on incorporating external evidence such as blended survival[3] or recent Bayesian methods for incorporating longer-term disease registry data, population data or elicited judgements into the extrapolation[4] can add substantial value to improving the extrapolations of immature data.
- Therefore, the results of this study show that clearer guidelines on how to handle immature OS data in survival analyses should be outlined by NICE.
- This would reduce the uncertainty surrounding the statistical methodology that should be employed when dealing with immature OS in survival analyses, improving submission quality and reducing discrepancies between the EAG/ERG and the submission with regards to the survival methods employed.
- Additionally, it is of note that there is some lack of consensus on the definition of mature data, with 22.2% of those STAs included which claimed immature data having reached median OS, a commonly used definition of data maturity.

REFERENCES

1. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. *Sheffield Report Decision Support Unit*. 2011;(0). www.nicedsu.org.uk/NICE DSU TSD Survival analysis. Updated March 2013.v2.pdf.
2. Rutherford M, Lambert P, Sweeting M, Pennington R, Crowther MJ, Abrams KR. NICE DSU Technical Support Document 21: Flexible methods for survival analysis. *Leic UK Dep Health Sci Univ Leic*. Published online 2020:1-97.
3. Che Z, Green N, Baio G. Blended Survival Curves: A New Approach to Extrapolation for Time-to-Event Outcomes from Clinical Trials in Health Technology Assessment. *Med Decis Making*. 2023;43(3):299-310. doi:10.1177/0272989X221134545.
4. Jackson C. survextrap: a package for flexible and transparent survival extrapolation. Published online June 6, 2023. doi:10.48550/arXiv.2306.03957.

Presented at ISPOR Europe 2023