Methods for Selecting Survival Extrapolations in Practice: A Systematic Review of Health Technology Assessments

Bütepage G[1], Haugli-Stephens T[2], Nanda S[3], Kaushik P[3], Lauppe R[3]

AstraZeneca, Stockholm, Sweden[1], AstraZeneca, Oslo, Norway[2], Quantify Research, Stockholm, Sweden[3]

MSR202

Background

- Extrapolation of time-to-event (survival) outcomes is commonly incorporated into economic evaluations for health technology assessment (HTA) where survival data are not fully mature.
- Projecting future survival outcomes is subject to uncertainty, requiring statistical and clinical assumptions and can impact health economic results.
- Different methods exist, such as standard parametric models (the hazard is assumed to follow a statistical distribution), or more flexible forms incorporating piecewise functions or splines models, among others.
- Guidelines from HTA authorities (1-4) outline steps for identifying appropriate models, but these also reference loosely defined concepts like clinical plausibility and validation with external data.
- Despite the availability of guidelines on methods, the selection of the extrapolation model continues to remain a major point of contention in HTA.

Objectives

The aim was to scrutinize the rationale and validation employed by marketing authorization holders (MAHs) and HTA agencies in selecting survival methods and parametric curves. Our objective was to ascertain whether established practices align with the parameters delineated in the guidelines, or if there are potential gaps in methodological guidelines where there are consistent established practices.

Methods

A systematic review of HTA reports was conducted to identify choices of extrapolation methods and the rationales behind those choices. HTAs within oncology and CVRM (cardiovascular, renal and metabolic disorders) which included information on the methods used for survival analysis were selected for assessment (**Figure 1**).

Up to five years of reports were reviewed from the following HTA agencies:

- NICE (England) from Jan 2019 to Dec 2023
- TLV (Sweden) from Jan 2019 to Dec 2023
- DMC (Denmark) from Jan 2021 (when cost-effectiveness was included in assessment) to Dec 2023
- NoMA (Norway) from Jan 2019 to Dec 2023

Search & Screening

NICE, TLV, NOMA and DMC HTA websites were searched for relevant submissions which were further screened for inclusion based on the presence of extractable outcomes on the survival extrapolation methodologies.

Data Extraction

- General submission details, survival extrapolation methods used and rationale for selection.
- Acceptance/rejection of MAH base case method by HTA bodies and rationale in case of disagreement.

Analysis

Data was assessed qualitatively and quantitatively to understand how MAH and HTA bodes justified and validated survival extrapolations.

Figure 1: Overview of methods for systematic review

Results

- A total of 272 HTA reports (NICE: 129, TLV: 58, DMC: 44, and NoMA: 41) were included in the review. The vast majority of HTA reports (94%) included were for oncological therapies, with most for lung cancer (16%) or breast cancer (14%).
- In both the MAH and HTA bodies, base case standard parametric models were most widely used, followed by piecewise approach and spline models.

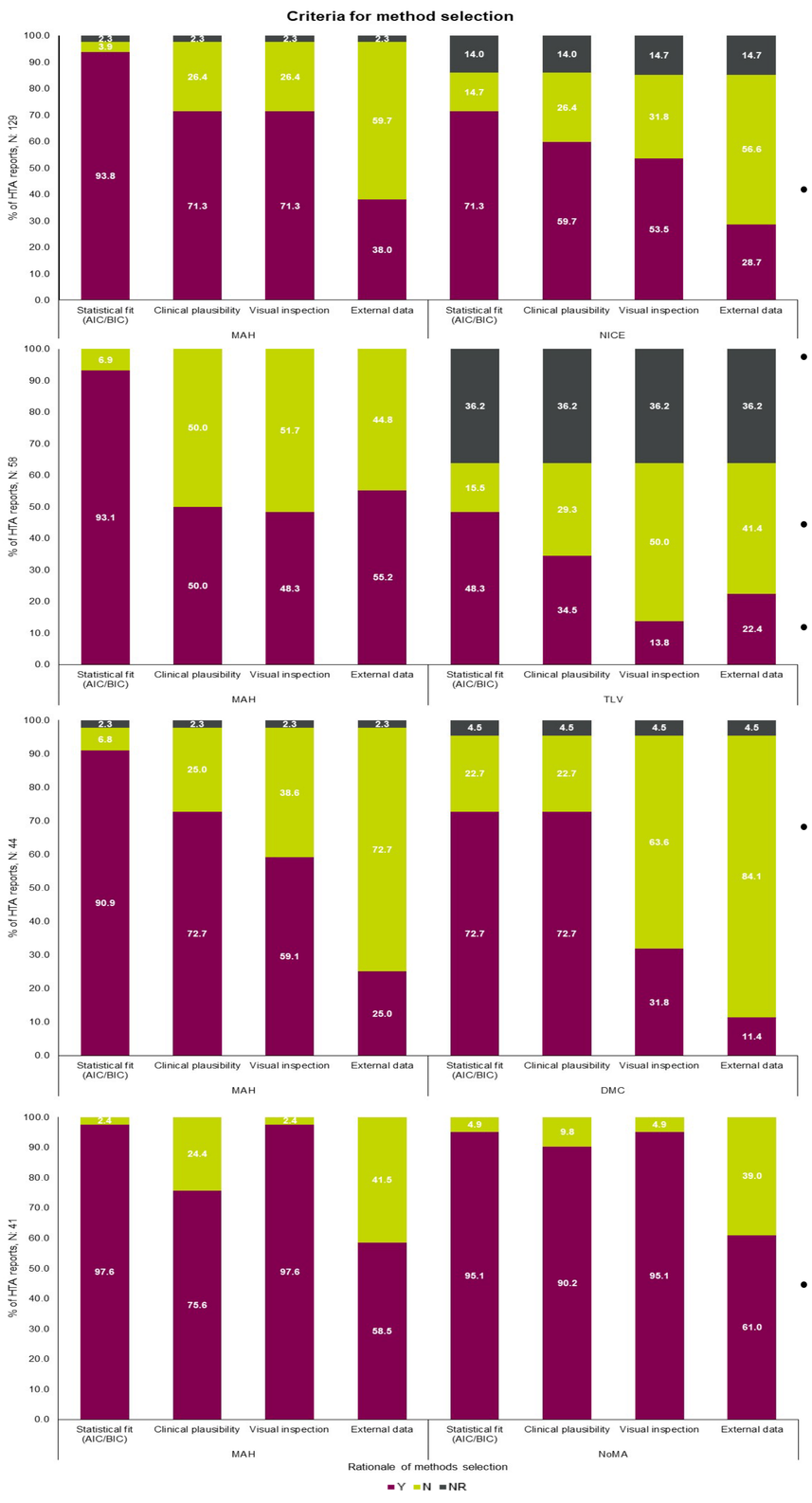

Figure 2: Criteria used for method selection by the MAH and HTA body

- Both MAH and HTA bodies cited guideline recommendations for selecting preferred curves (e.g., statistical and visual fit, fit to external data, or clinical plausibility). In most cases, MAH and HTA bodies referred to more than one of the rationales to support their choices (**Figure 2**).
- Fit to trial data (statistical or visual) was most commonly given as a rationale for curve selection by MAH (>93%) and was referred to most commonly by HTA bodies for their base case preferences (48% for TLV – 98% for NoMA).
- Reports referenced MAHs claims on clinical plausibility in 50-75% of cases, depending on the country, but clinical plausibility was discussed almost as often as statistical fit by HTA bodies (35% for TLV – 90% for NoMA).
- External data was cited by MAH in 25-55% of cases but was often referenced less by HTA bodies (11% for DMC – 61% for NoMA).
- In case of disagreement with the MAH, HTA bodies generally cited multiple reasons for changes in the base case extrapolation, mostly commonly statistical fit and clinical plausibility, though external data was also cited in many cases.
- Clinical plausibility was in most cases based on input from clinical experts. NICE, TLV and NoMA consulted clinicians with certain questions, typically only a limited number of clinicians per case, while DMC made use of the expert committee created as part of the assessment. Inputs from clinical experts focused on validation of survival estimates at landmark points, the plausibility and timing of a cure assumption or plateau in the extrapolation, assessment of the hazard plot and the comparability of study population with local population. Clinical plausibility (however defined) was often the principal factor in HTA body curve selection.
- External data was used in numerous ways to inform and validate curve selections, but in many cases, it was used for comparison of survival at landmark points. The acceptance of the external data by HTA agencies largely depended on the fit for purpose of the external data, e.g., how well the patient population in the external data matched the indication and the trial/local population.

Conclusions and Recommendations

Statistical fit remains the most cited rationale for curve selection, despite recent guidance that it is one of the least important criteria (5).	Statistical goodness of fit should primarily be used for excluding curves with bad fit to trial data, as it can be of limited use in determining the quality of long-term extrapolations.
Establishing clinical plausibility is essential but often based on clinician interviews and is not consistently operationalised across assessments.	MAH and HTA bodies should focus on obtaining transparent and generalisable structured clinical input. Guidance on systematic methods for collecting expert input for survival extrapolation could be beneficial.
External data is less frequently used even though recommended in guidelines. This could be due to lack of appropriate, reliable data and concerns around the fit to local populations.	Identifying or generating relevant external data to support survival extrapolation that matches the decision problem should be prioritised. Research should focus on how to use external evidence to inform survival extrapolations for novel interventions where, for example, longer-term follow-up in the same disease and treatment line is not available.
HTA bodies often reject non-standard parametric extrapolation methods (e.g., piecewise, spline, mixture cure), believing that a standard parametric model suffices.	If it is believed that there are multiple turning points in the hazard, requiring non-standard models, the relevance of including such turning points must be clearly demonstrated based on external evidence and/or structured clinical input, where data are immature.
In several submissions, the criteria for method selection were not clearly reported, particularly from HTA bodies.	As survival extrapolation can significantly impact cost-effectiveness, the assumptions behind the selection of survival models by both MAH and HTA bodies should be transparently reported in public documentation.

Abbreviations

AIC: Akaike information criteria, BIC: Bayesian information criteria; DMC: Danish Medicines Council (Medicinrådet); HTA: Health technology assessment; MAH: Marketing authorization holder; NoMA: Norwegian Medicinal Products Agency (Direktoratet for medisinke producter), NICE: National Institute for Health and Care Excellence, TLV: Tandvårds- och läkemedelsförmånsverket (Dental and Pharmaceutical Benefits Agency)

References

1. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2011.
2. Rutherford MJ, Lambert PC, Sweeting MJ, Pennington R, Crowther MJ, Abrams KR, Latimer NR. NICE DSU Technical Support Document 21: Flexible Methods for Survival Analysis. 2020.
3. Medicinrådet. Anvendelse af forløbsdata i sundhedsøkonomiske analyser (Version 1.1). 2020.
4. Norwegian Medicinal Products Agency. Submission guidelines for single technology assessment of medicinal products. 2024.
5. Palmer S, Borget I, Friede T, Husereau D, Karnon J, Kearns B, et al. A Guide to Selecting Flexible Survival Models to Inform Economic Evaluations of Cancer Immunotherapies. Value Health. 2023;26(2):185-192.

Supported by AstraZeneca

Presented at ISPOR Europe 2024, Barcelona, Spain, 17-20 November 2024