POSTER# MSR62

Keeping Up With the Times: Emerging Methods for More Precise Survival Estimates to Inform Cost-Effectiveness Analysis

Authors: Liu, YH.,[1] Gulas, I.,[2] Tyas, E.,[1] Tan, YZ.,[3] Chang JYA.,[1] Kearns, B.,[1] Westley, T.[4]

Affiliations: 1. Lumanity, Sheffield, England, UK; 2. Lumanity, Toronto, Canada; 3. Lumanity, Utrecht, The Netherlands; 4. Lumanity, Bethesda, MD, USA

Acknowledgments: The authors thank Su Yi Cheng, Lumanity, Sheffield, for technical support with the model build.

BACKGROUND

- Emerging flexible time-varying network meta-analysis (NMA) methods can more precisely estimate the survival improvements of novel therapies compared with existing treatments[1, 2]
- Precedents in the literature often highlight advancements in immune-oncology, with novel biological mechanisms of action requiring more flexible models to maximally capture long-term survival benefits and provide plausible extrapolations[3]
- Time-varying NMA models can mitigate statistical limitations when the proportional hazard (PH) assumption for comparing survival curves is violated
- However, these advanced methods remain underutilized in North American regulatory and reimbursement environments compared with UK and EU practice, despite their potential to maximally capture long-term benefits of a new therapy.[4,5] Also, it is recognized that health technology assessment (HTA) guidance is currently limited for time-varying NMA in terms of model selection, NMA implementation in cost-effectiveness models (CEMs), and the assessment of impact of the NMA estimates on CEM results

OBJECTIVES

- We aimed to demonstrate how application of various time-varying NMA methods can impact final estimates in a treatment cost-effectiveness analysis (CEA), representing these differences with associated incremental cost-effectiveness ratios (ICERs)
- The goal was to provide transparency in discussing the complexities of applying these emerging, time-varying NMA methods, sharing key technical and clinical assumptions, advantages, and limitations

METHODS

- We revisited a recent multiple technology appraisal (ID3760) from the UK's National Institute for Health and Care Excellence (NICE) in advanced renal cell carcinoma, where PH violations for progression-free survival (PFS) and overall survival (OS) curves were previously identified.[6] The appraisers chose not to apply time-varying NMA results in the CEA, stating difficulty in interpretating their unintuitive results. NICE submissions are renowned for their high-quality reporting and methodological rigor, positioning NICE as a leader in advancing HTA methodologies[7]
- We applied available, updated clinical trial data to the NICE base case, while creating different CEA scenarios by only changing the type of NMA model estimating PFS and OS hazard ratios (HRs). Time-varying NMA methods included multivariate parametric, fractional polynomial (FP), restricted cubic spline (RCS), and piecewise exponential (PWE) NMA. Resultant ICERs were compared with CEA estimates using constant (Cox PH) NMA.

Figure 1. ID3760: Network diagram for the intermediate/poor risk subgroup (PFS and OS)

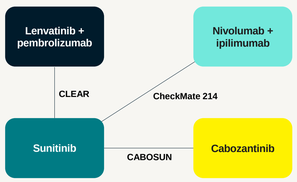


Table 1. Study Scope from 2022 NICE MTA (TA858, ID3760)[6]

Population	Untreated advanced RCC: Intermediate/poor risk subgroup (IMDC criteria)
Interventions	• Lenvatinib + pembrolizumab (CLEAR)[8]
Comparator*	• Cabozantinib (CABOSUN)[9] • Nivolumab + ipilimumab (CheckMate 214)[10]
Outcomes	• Progression-free survival (PFS) • Overall Survival (OS)
Setting:CEA	• Incremental cost per QALY • NHS and PSS perspective • Three-state PSM

Key: IMDC, International Metastatic Renal Cell Carcinoma; MTA, multiple technology assessment; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; NMA, network meta-analyses; OS, overall survival; PFS, progression-free survival; PSS, Personal Social Services; PSM, partitioned survival model; QALY, quality-adjusted life year; RCC, renal cell carcinoma.

Note: *Sunitinib is used as a common comparator to connect the network but would not be a comparator of interest for the intermediate/poor risk subgroup (see Figure 1)

Network meta-analysis

- Patient-level data for intermediate to poor risk groups were reconstructed from published PFS and OS KM graphs[11,12]
- Given the limited evidence base with only one trial per treatment comparison, a Bayesian NMA framework using fixed effects models was applied[12-14]
- Best fitting NMA models were determined based on NICE recommended criteria for survival extrapolation, including visual fit to KM graphs, statistical parsimony (e.g. AIC, BIC, DIC), biological plausibility as discussed in ID3760, and convergence diagnostics[15]

Multivariate parametric NMA (2-step)[16]

- Seven parametric distributions were used for parametric NMA (exponential, Weibull, gamma, Gompertz, generalized gamma, log-normal, log-logistic). Parametric NMA requires the same survival distribution to be fitted for all treatment arms across the network. The accuracy of this model hinges on the chosen distribution: a mis-specified functional form may restrict the hazard shape and fail to capture complex features seen in smoothed hazard plots

Fractional polynomial (FP) NMA[17]

- First and second order FP NMA models were fitted to the data. The choice of powers included all combinations of {-2, -1, -0.5, 0, 0.5, 1, 2}, giving a total of 35 models (7 and 28 for first and second order FP, respectively). All treatments are assumed to follow the same polynomial functional form (within the same set of pre-specified power terms) for their baseline hazard, with differences captured by treatment-specific coefficients

Restricted cubic spline (RCS) NMA[18]

- 1-, 2-, and 3-knot RCS NMA models were fitted using equally distributed knots specific to each treatment arm based on the log of uncensored survival times. RCS methods assume that the log cumulative hazard function (i.e. the baseline hazard of each trial) can be appropriately and smoothly modeled with splines across time, and any non-PH can be captured through time interactions

Piecewise exponential (PWE) NMA[19]

- PWE NMA assumes treatment effects for each of the time intervals are represented by constant HRs, following exponential distributions. Treatment effects are assumed to be constant on the hazard scale within each time interval. Three PWE models were used with different time breaks: (1) 6 months, (2) 12 months, (3) 6 and 12 months

Constant NMA[20]

- Time-constant semi-parametric (Cox PH) NMA models assume that the hazard rate is proportional over time. For the purpose of this exercise, constant NMA results were used for comparison

Cost-effectiveness analysis

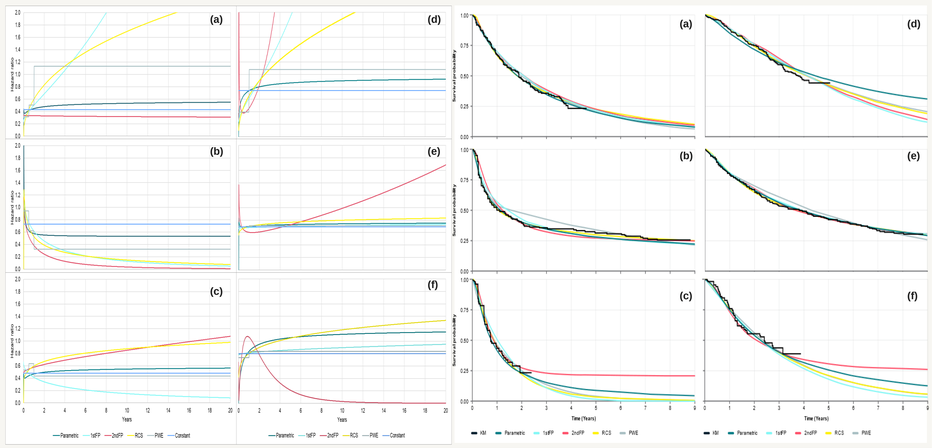
- To reflect TA858, a three-state partitioned survival model was constructed. All cost, healthcare resource use (HCRU), and utility inputs from TA858 were applied in the model, with costs inflated to 2023/24. Time to treatment discontinuation was assumed to be equal to PFS due to data unavailability. As sunitinib was the common node in the network, when comparator time-varying HRs were generated from an NMA method, these were applied onto the modeled sunitinib arm from CLEAR to derive survival estimates over time. For sunitinib the log-normal and generalized gamma models were used for OS and PFS, respectively, due to a combination of their good visual fit and clinical plausibility
- Time-varying HRs were derived from the best fitting models from each of the various NMA results:
  - For parametric NMA, survival estimates for all three comparators and pooled sunitinib were generated directly from NMA parameters, and the hazards were compared in a pairwise fashion for each comparator against sunitinib over time
  - For RCS NMAs, survival estimates for all three comparators were generated directly from the NMA parameters
  - For FP NMAs, computed from d0, d1, and d2 parameters using equation from Jansen 2011[17]
  - For PWE NMAs, obtained directly as statistical outputs
- Costs and QALYs were summed across a 40-year horizon for each comparator with a 3.5% discount. Pairwise comparisons were conducted between L+P against N+I and cabozantinib where incremental costs and QALYs were computed. The ICER was then estimated by dividing the incremental costs by the incremental QALYs

RESULTS

Network meta-analysis

Figure 2. HR from NMA (best fitted models)* (a) PFS L+P (b) PFS N+I (c) PFS cab (d) OS L+P (e) OS N+I (f) OS cab

Figure 3. Survival estimates from NMA overlayed with KM curves (best fitted models)* (a) PFS L+P (b) PFS N+I (c) PFS cab (d) OS L+P (e) OS N+I (f) OS cab



Cost-effectiveness analysis

- Table 2 and Table 3 show resultant ICERs from the best fitted and alternative (second best selection) models from each NMA method, respectively

Table 2. Incremental cost-effectiveness results versus L+P based on various NMA approaches (best fitted models)

NMA model/Treatment	versus Nivolumab + ipilimumab			versus Cabozantinib		
	Incremental cost	Incremental QALY	ICER	Incremental cost	Incremental QALY	ICER
Constant	£106,764	0.334	£319,332	-£4,375	0.560	L+P dominates
PWE *	£115,023	-0.896	L+P dominated	£10,463	-0.465	L+P dominated
RCS*	£99,313	-1.172	L+P dominated	£33,776	-0.268	L+P dominated
First-order FP*	£79,907	-1.915	L+P dominated	-£32,442	-1.366	L+P is inferior
Second-order FP*	£70,624	-2.249	L+P dominated	-£12,777	-3.208	L+P is inferior
Parametric*	£122,739	0.015	£8,442,537	£38,624	0.823	£46,933

Table 3. Incremental cost-effectiveness results versus L+P based on various NMA approaches (alternative models)

NMA model/Treatment	Versus Nivolumab + ipilimumab			versus Cabozantinib		
	Incremental cost	Incremental QALY	ICER	Incremental cost	Incremental QALY	ICER
PWE**	£114,615	-0.830	L+P dominated	£10,485	-0.366	L+P dominated
RCS**	£99,602	-1.175	L+P dominated	£30,324	-0.321	L+P dominated
Parametric**	£120,393	-0.568	L+P dominated	£63,799	0.389	£164,123

Note: *Best fitted models for OS: piecewise (time break = 12 months); RCS (1 knot); first-order FP (P1=0.5); second-order FP (P1=0.5, P2=0); parametric (log-normal). For PFS: piecewise (time break= 6, 12 months); RCS (2 knots); first-order FP (P1-0.5); second-order FP (P1=0.5, P2=0); parametric (generalized gamma).

**Alternative fitted models for OS: piecewise (time break = 6, 12 months); RCS (2 knots); parametric (log-logistic). For PFS: piecewise (time break= 12 months); RCS (3 knots); parametric (log-normal). For FP NMA, there is no recommendation on alternative models due to the convergence issue and implausibility of the extrapolations.

LIMITATIONS

- **NMA model selection:**
  - Nuanced decisions for each NMA type, such as powers and time intervals for FP NMA, knot placement for RCS NMA, and time breaks for PWE NMA, introduced uncertainty in the decision making. Different model approaches are not easily comparable, and clinical judgment of these additional models outside of the precedent NICE decision problem was not currently feasible
  - Forty-eight time-varying NMA models were analyzed for each endpoint, requiring significant computational effort. Testing the best fitting FP and parametric NMAs was time-consuming due to numerous, reasonable options for possible model inputs to be tested. Convergence was challenging for FP NMA, especially for 2nd order models. PWE NMAs were simpler to implement but may lack clinical plausibility due to step-wise HRs
- **Early-cycle effect:** Implausibly large HRs from some models led to zero survival estimates in the CEA for the first model cycle. Therefore, no treatment effect was applied in the first month to remove implausible estimates
- **Anchoring survival:** CEA typically relies on pivotal trials with IPD, while ITC estimates for HTAs use aggregate data from comparator trials. In this study, NMA-derived HRs were anchored on parametric survival extrapolation of sunitinib from the CLEAR trial. This CEA approach requires varied considerations for deriving anchored survival estimates
- **Uncertainty analysis:** The only scenario analyses conducted were to assess uncertainty in selection of the NMA model. In the future, it would be beneficial to conduct probabilistic sensitivity analysis to explore different settings of variance of the NMA models. It would also be interesting to quantify the uncertainty across different models and its resulting impact on ICERs and/or decision making[21]

CONCLUSIONS

- Recognizing that the decision to select the most appropriate NMA model for estimating comparative efficacy occurs well before generating ICER results from various models, this exercise helps to inform the broad range of final CEA conclusions where ICERs are sensitive to the selected NMA method
- It was found that while keeping all other CEA input data equal, across the NMA models, the derived relative treatment effects and the resultant ICERS were very different and informed a broad range of possible cost-effectiveness interpretations
- Considering the limitations described above, it is essential to evaluate the advantages and disadvantages unique to each NMA type. This assessment should consider the core elements of the analysis, such as the validity of the PH assumption, the shape of the trial KM curves and hazard plots, the likely user's time investment, acceptability to decision makers and the accessibility of clinical expertise

REFERENCES

1. Heeg et al. J Comp Eff Res. 2022;11(15):1121–1133. 2. Rutherford et al. NICE DSU TSD 21. 2020. 3.Palmer et al. Value Health. 2023; 26(2):185–192. 4. Aiello et al. Value Health. 2025a; 28 (S1). 5. Aiello et al. Value Health. 2025b; 28 (S1). 6. Fleeman et al. Univ Liverpool report. 2018. Data on file. 7. Kumar et al. Int J Technol Assess Health Care. 2025;41(1):e21. 8.Motzer et al. J Clin Oncol. 2024; 42(11):1222–1228. 9.Choueiri et al. Eur J Cancer. 2018; 94:115–125. 10.Tannir et al. Ann Oncol. 2024; 35(11):1026–1038. 11.Liu et al. BMC Med Res Methodol. 2021; 21:111. 12. Dias et al. NICE DSU TSD 2. 2011. 13. Cope et al. Value Health. 2023; 26(4):465–476. 14. Sturtz et al. J Stat Softw. 2005;12(3):1–16. 15.Latimer. NICE DSU TSD 14. 2011. 16. Cope et al. Res Synth Methods. 2020;11(3):443–456. 17.Jansen. BMC Med Res Methodol. 2011;11:61. 18. Freeman & Carpenter. Res Synth Methods. 2017;8(4):451–464. 19. Freeman et al. Stat Methods Med Res. 2022; 31(5):839–861. 20.Cox. J R Stat Soc B. 1972;34(2):187–202. 21. Kearns et al. PharmacoEconomics 38, 193–204 (2020).

An electronic version of the poster can be viewed by scanning the QR code.

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; Cab, cabozantinib; DIC, deviance information criterion; ICER, incremental cost-effectiveness ratio; FP, fractional polynomial; HTA; health technology assessment; LY, life year; L+P, lenvatinib + pembrolizumab; MTA, multiple technology assessment; NICE, National Institute for Health and Care Excellence; NMAs, network meta-analyses; N+I, nivolumab + ipilimumab; OS, overall survival; PFS, progression-free survival; PH, proportional hazard; PWE, piecewise exponential; QALY, quality-adjusted life year; RCS, restricted cubic spline.