Comparing Alternative Extrapolation Methods Using Standard Partitioned Survival Model Functionality in the Presence of Converging Survival Data: A Case Study in Renal Cell Carcinoma

EE141

Paulina Kazmierska,[1] Neda Aminnejad,[2] George Bungey[1]

[1]Thermo Fisher Scientific, London, UK; [2]Thermo Fisher Scientific, Toronto, Canada

Background

- In specific circumstances, combinations of parametric survival extrapolations may produce implausible crossings between two comparators due to the characteristics of the underlying Kaplan-Meier (KM) data.
- This was observed with the CLEAR trial[1] data in NICE TA858[4] when comparing lenalidomide in combination with pembrolizumab (LEN+PEM) versus sunitinib (SUN) in patients with renal cell carcinoma (RCC).
- In the August 2020 cut of CLEAR trial data, while LEN+PEM showed improved overall survival (OS) in the short term, crossing occurred between the KM curves at approximately 33 months.

Objectives

- **The objective of this case study was to explore the impact of combining typical functionalities (treatment effect waning, KM+parametric fit) included in partitioned survival models (PSMs) with standard parametric models as alternative extrapolation approaches, compared to conservatively assuming equivalence at the crossing point or unadjusted extrapolation, considering multiple OS data cuts from the CLEAR trial.**

Methods

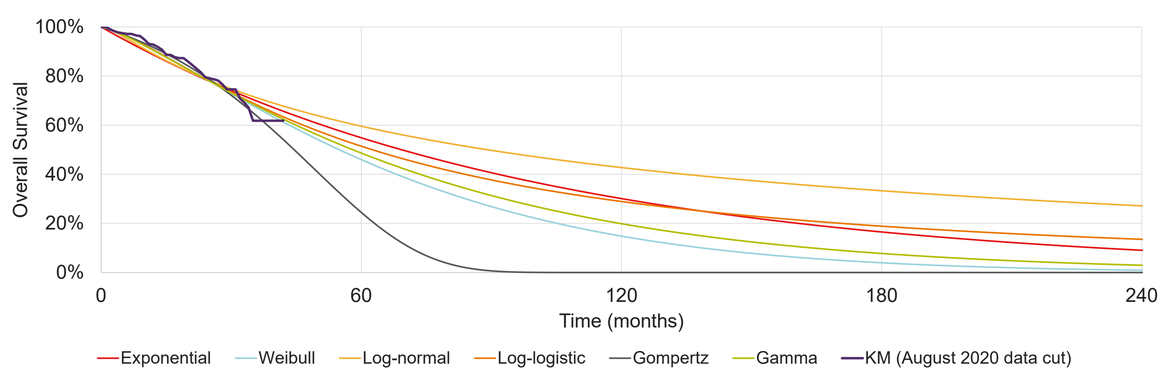
- Published OS data from the CLEAR trial were digitized using WebPlotDigitizer software[2] and pseudo-individual patient data generated by Guyot Algorithm.[3] A series of standard parametric distributions was then fitted in the R flexsurv package to the August 2020 data cut.
- Six standard parametric functions were then fitted (exponential, Weibull, lognormal, loglogistic, Gompertz, generalized gamma). Models were selected based on statistical, visual fit, comparisons of hazard profiles for the parametric models against smoothed hazard plots for the KM data, and on the basis of plausibility of long-term predictions according to UK clinical expert expectations for SUN from technology appraisal (TA) 858 by the National Institute for Health and Care Excellence (NICE; <20% at 10 years).[4] In line with NICE decision support unit (DSU) technical support document (TSD) 14 guidance,[5] the same type of parametric model was selected for both comparators in the absence of a strong rationale to support different types.
- Given the crossing of OS KM curves observed in the CLEAR trial and potential uncertainty around long-term extrapolations for OS, the following approaches were also explored: (1) assuming equivalence at the crossing point; (2) assuming equivalent efficacy to SUN for LEN+PEM at the start of convergence between the two treatments, after which LEN+PEM OS hazards are set equal to SUN; (3) exploring a combined KM+parametric extrapolation approach using a truncated KM curve.
- Long-term extrapolations were then visually compared with the final data cut KM curve (July 2022).[6]

Results

Unadjusted Extrapolations

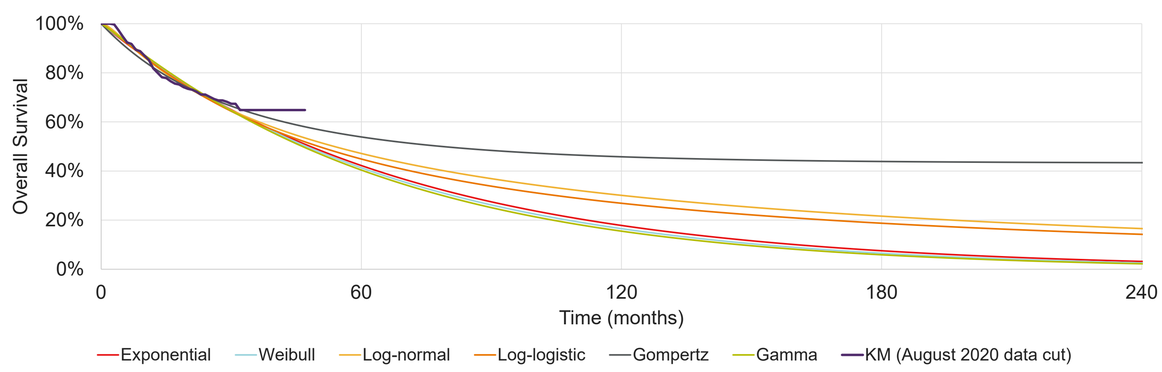
- Joint parametric distributions were not suitable due to curve crossings on the log-cumulative hazard plot with clearly non-parallel hazard plots, as well as the result of the formal assessment of the proportional hazards (PH) assumption via the Schoenfeld residuals test (P<0.0001).
- Among individual fits, the Gompertz and log-normal distributions produced the best statistical fit according to both Akaike information criterion (AIC) and Bayesian information criterion (BIC) for LEN+PEM and SUN, respectively. Most SUN fits showed relatively poor visual fit to the observed data, and almost all produced aggressive curve crossing (**Figure 1** and **Figure 2**). The exponential models, while producing poor visual fits, were considered the most plausible set of single-fit distributions given clinical expert expectations for long-term OS discussed in NICE TA858[4] (<20% for patients starting treatment with SUN) without curve crossing and were selected for both comparators in line with NICE DSU TSD14 guidance[5].
- Among the remaining individual fits, the Weibull model produced the least sharp curve crossing and met clinical expert expectations from NICE TA858[4] for long-term OS in the SUN arm. This model was explored for both comparators in the equivalence assumption approach.

Figure 1. Long-term Single Parametric OS Predictions for LEN+PEM Using August 2020 Data Cut



Abbreviations: KM = Kaplan-Meier; LEN+PEM = lenalidomide in combination with pembrolizumab; OS = overall survival

Figure 2. Long-term Single Parametric OS Predictions for SUN Using August 2020 Data Cut



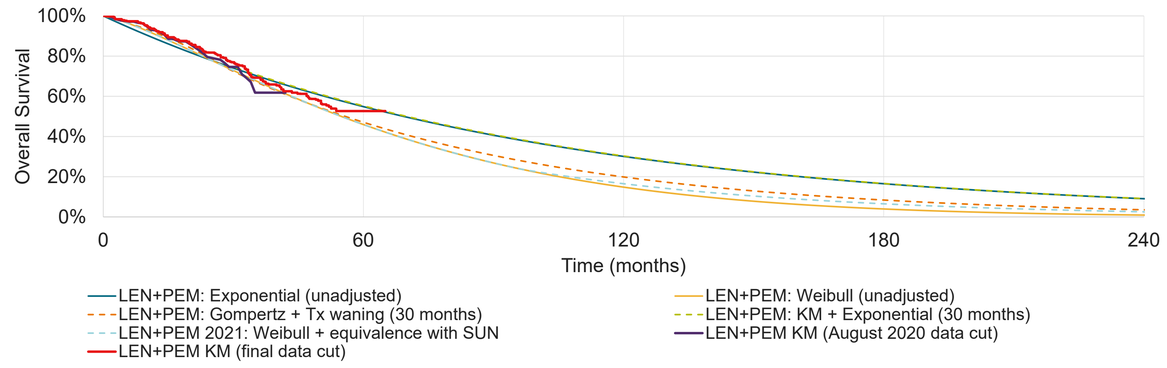
Abbreviations: KM = Kaplan-Meier; LEN+PEM = lenalidomide in combination with pembrolizumab; OS = overall survival

Results (cont.)

Method 1: Equivalence Assumption

- Single Weibull models were applied in combination with the SUN equivalency assumption mechanic to allow for a smoother convergence between LEN+PEM and SUN curves without actual crossing. This approach helped avoid overpredicting the tail of the original KM curve; however, it resulted in more conservative extrapolation for LEN+PEM compared with the unadjusted individual exponential fit and underprediction of the tail of the final KM curve.

Figure 3. Approaches for Long-term OS Predictions for LEN+PEM



Abbreviations: KM = Kaplan-Meier; LEN+PEM = lenalidomide in combination with pembrolizumab; OS = overall survival; SUN = sunitinib; Tx = treatment
In Methods 2 and 3, two switch points were explored (24 and 30 weeks); as curves are closely overlapping, only 30-week switch point data is visualized for both methods

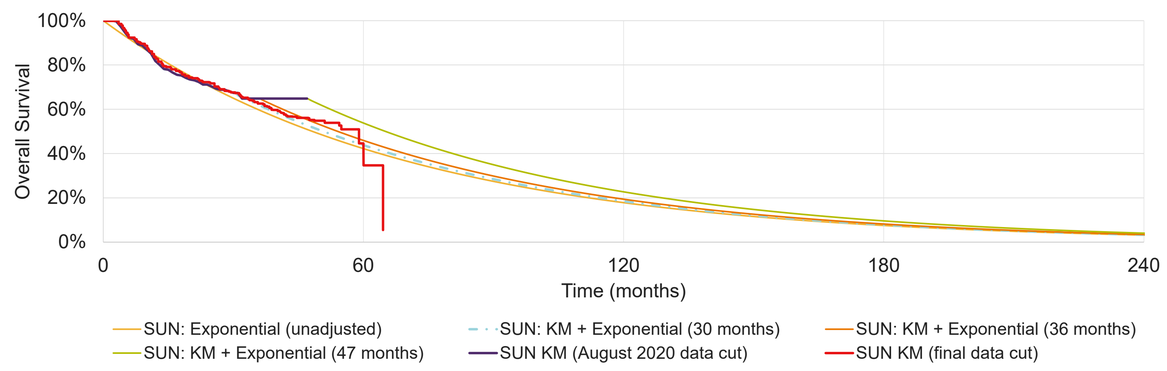
Method 2: Treatment Effect Waning

- The treatment effect waning approach for LEN+PEM explored two switch points: (1) at 24 and (2) at 30 weeks, after which the OS hazard of LEN+PEM was set equal to that of SUN. The switch points were selected based on visual inspection of the KM curves; at these times, curves appeared to begin converging. For LEN+PEM the Gompertz model, the best statistical fit according to both AIC and BIC, was used.
- Both switch points produced nearly overlapping sets of curves. In both cases, the treatment effect waning approach helped avoid overpredicting the tail of the original KM curve, similar to Method 1, while producing a conservative extrapolation for LEN+PEM compared with the unadjusted individual exponential fit with underprediction of the tail of the final curve.

Method 3: KM+Parametric Extrapolation

- The KM+parametric extrapolation approach for LEN+PEM used an individual exponential fit and explored two switch points: (1) at 24 and (2) at 30 weeks; at these times, curves appeared to begin converging on the original KM plot. Both switch points produced nearly overlapping sets of curves. Both approaches resulted in a slightly better visual fit to the observed data to the original unadjusted exponential fit. Compared with the KM curves, the KM+parametric extrapolation approach produced an overprediction of the original 2020 data cut KM data but produced a close fit to the tail of the final 2022 data cut KM curve.
- The KM+parametric extrapolation approach for SUN used an individual exponential fit and explored three switch points: (1) at 47 months, corresponding to the maximum follow-up for the SUN arm at the August 2020 data cut off; (2) at 36 months, when the number of patients at risk dropped below 10; and (3) at 30 months, shortly before the original curves crossed. Switching at 47 months resulted in a clinically implausible estimate (>20% at 10 years), whereas the other two switch points resulted in slightly more optimistic extrapolations compared to the original unadjusted exponential fit.

Figure 4. Approaches for Long-term OS Predictions for SUN



Abbreviations: KM = Kaplan-Meier; SUN = sunitinib

Conclusions

- **The results of our case study show that relatively straightforward functionalities commonly implemented in PSMs may provide reasonable alternative extrapolations to support further scenario analyses in the presence of clinically implausible converging survival extrapolations between study arms, especially in cases where proportional hazards are clearly violated and joint parametric fits appear inappropriate, as seen for the CLEAR trial data.**
- **However, careful rationalization of appropriate switch points for treatment effect waning and KM+parametric approaches is required.**

References

1. Motzer R, et al. *N Engl J Med*. 2021;384(14):1289-300.
2. WebPlotDigitizer, version 5.2. https://automeris.io
3. Guyot P, et al. *BMC Med Res Methodol*. 2012;12:9.
4. National Institute for Health and Care Excellence (NICE). TA858: Lenvatinib with pembrolizumab for untreated advanced renal cell carcinoma. 2023. https://www.nice.org.uk/guidance/ta858
5. Latimer NR. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2011. https://www.ncbi.nlm.nih.gov/books/NBK395885/pdf/Bookshelf_NBK395885.pdf
6. Motzer RJ, et al. *J Clin Oncol*. 2024;42(11):1222-8.

Disclosures

All authors are employees PPD™ Evidera™ Health Economics & Market Access, Thermo Fisher Scientific. The views expressed in this study are those of the authors and not necessarily those of Thermo Fisher Scientific. The authors do not have any conflicts of interest.

Acknowledgments

The authors would like to thank Caroline Cole and Kawthar Nakayima of Thermo Fisher Scientific for their editorial and graphic contributions.

Presented at ISPOR —The Professional Society for Health Economics and Outcomes Research—Europe | 9–12 November 2025 | Glasgow, Scotland, UK