Parametric Mixture Models: An Advanced Approach to Address Complex Hazards in Survival Analysis

MSR82

Shubhram Pandey[1], Parampal Bajaj[1], Barinder Singh[2], Supreet Kaur[1], Akanksha Sharma[1]

[1]Heorlytics, Mohali, India; [2]Pharmacoevidence, London, UK

CONCLUSIONS

- This analysis involved combining of models since single distribution models fail to consider variation in hazard functions which can be used to predict the long-term outcomes in survival
- Log-logistic and Gompertz mixture models were considered as suitable for describing intricate survival curves, considering patient characteristics and giving accurate long-term treatment effects assessments
- This model flexibility is useful in the economic analysis where it is important to extrapolate data into the distant future to capture the complete value of a treatment

PLAIN LANGUAGE SUMMARY

- This study focused on improving how we predict the long-term effects of treatments on patient survival for health assessments. Instead of using a single model to fit survival data, we explored combining two different models to better capture complex survival patterns.
- By using a mix of different statistical models, we were able to better identify groups of patients who respond differently to treatments. This approach allowed us to more accurately reflect real-life survival patterns and potential benefits of treatments over time.
- Our results showed that combining two models (Log-logistic and Gompertz) worked best, giving a more accurate fit than using just one model alone.

INTRODUCTION

- In Health Technology Assessment (HTA), survival modeling is essential for estimating the benefits of treatments in terms of survival and quality of life when these effects are extrapolated to post trial periods to analyze the efficacy of the treatments
- Traditionally single-parametric distributions, such as Exponential or Weibull, is assumed that the hazard functions are constant, which are not appropriate for characterizing non-linear hazard patterns when patient population heterogeneity is taken into consideration
- Parametric mixture models address limitations of traditional models by combining multiple distributions, allowing for separate modeling of patient subgroups with varying risk characteristics
- This approach provides a probabilistic framework to estimate each patient's likelihood of belonging to different survival profiles, without definitively categorizing them, which supports complex hazard function modeling and enhances long-term prediction accuracy

OBJECTIVE

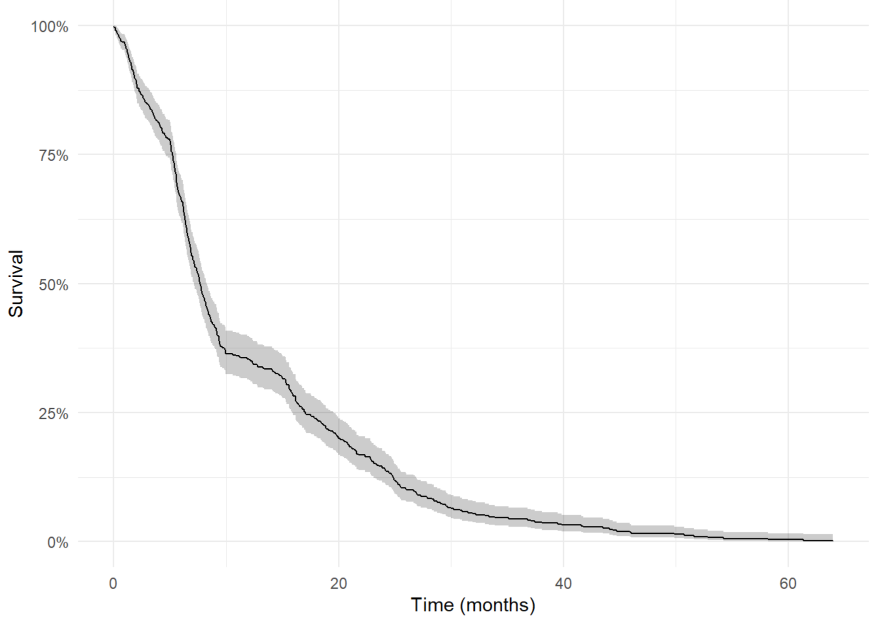
- To fit parametric mixture models on time-to-event data with complex hazard function i.e., hazard function is increasing at certain time point and then decreasing
- To assess the fit of parametric mixture models against single-distribution models using Akaike Information Criteria (AIC), Bayesian information criteria (BIC), and visual inspection, with the objective of predicting long-term outcomes.

METHODS

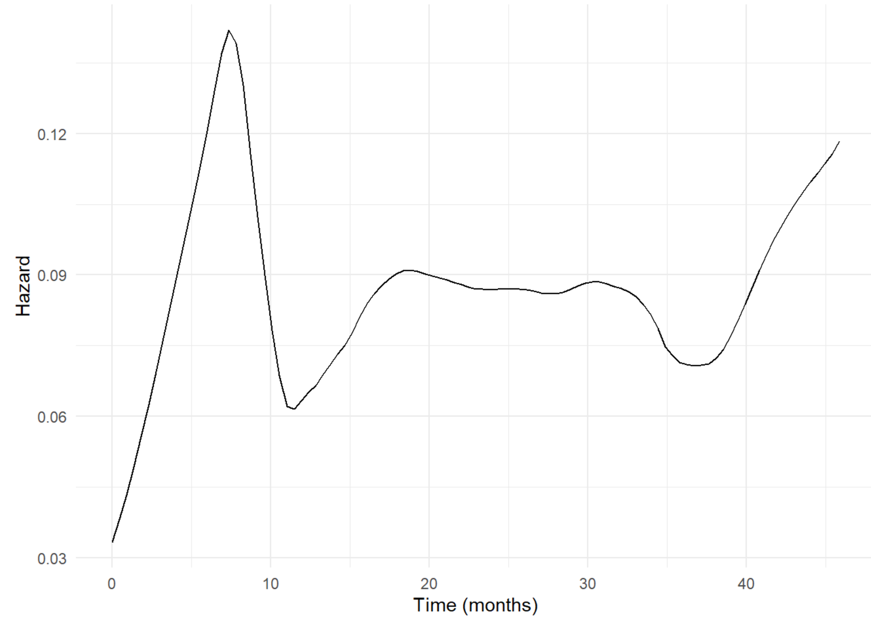
Step 1: Data Simulation

- Pseudo Individual patient-level data (IPD) was generated using the Guyot algorithm, allowing for detailed time-to-event analysis
- The simulated data included two distinct subpopulations (in proportion of 35% and 65%) with unique hazard trajectories, leading to a complex hazard function

Kaplan-Meier curve



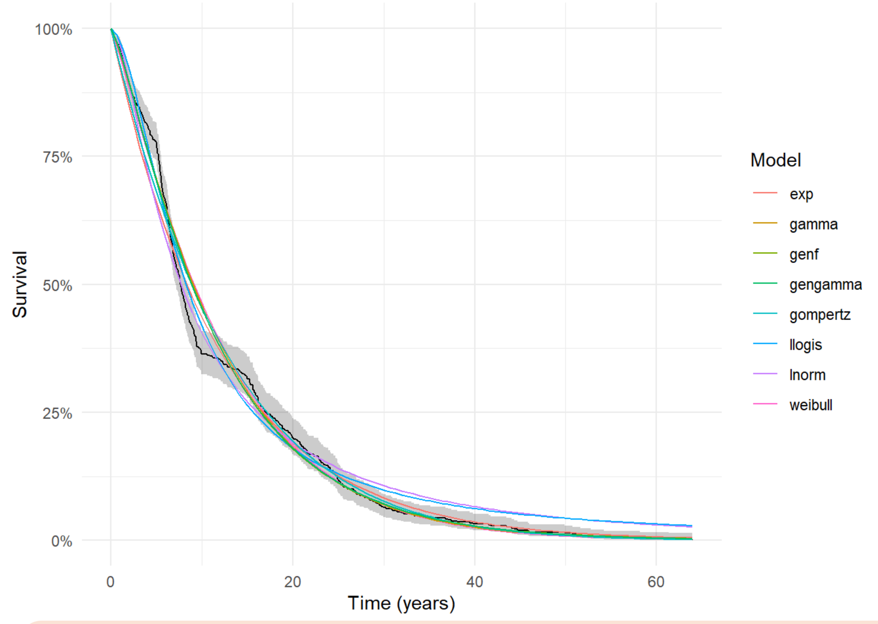
Hazard function curve



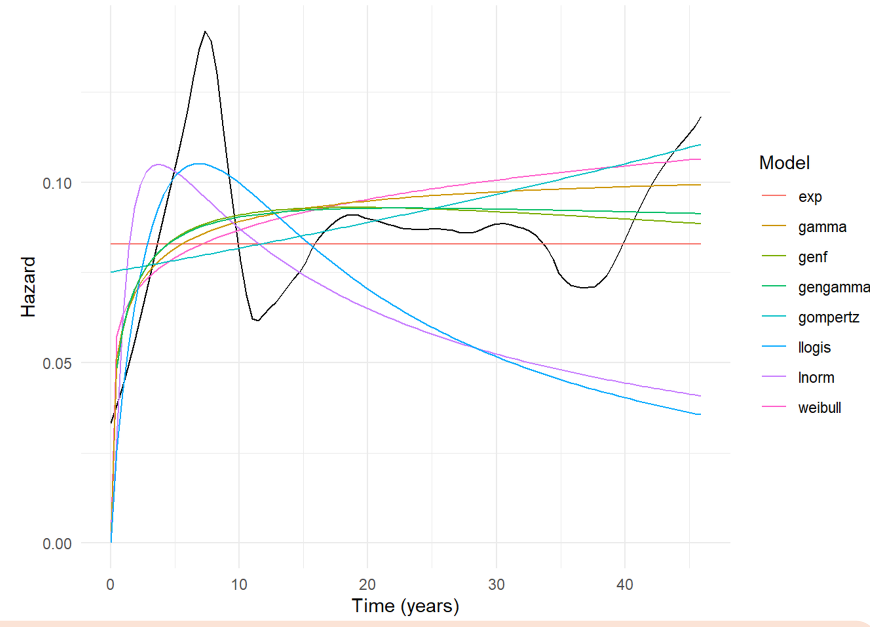
Step 2: Single distribution model fitting

- Standard parametric models, including Exponential, Weibull, Log-normal, Log-logistic, Gompertz, Gamma, and Generalized Gamma, were fitted to the time-to-event data using *flexsurvreg* function in R (v4.3.1)
- The fitted distributions were plotted against Kaplan-Meier and observed hazard function for visual inspection and AIC, BIC were presented for comparison of fitted models

Kaplan-Meier curve



Hazard function curve



The Kaplan-Meier and hazard plots indicate that none of the single parametric models adequately capture the complex hazard pattern over time, leading to either overestimation or underestimation of the observed Kaplan-Meier curve

Step 3: Parametric mixture models

- The mixture model is defined by the following equation:

$$\sum_{k=1}^{K} p_k\, f_k(t), \text{ where } \sum_{k=1}^{K} p_k = 1$$

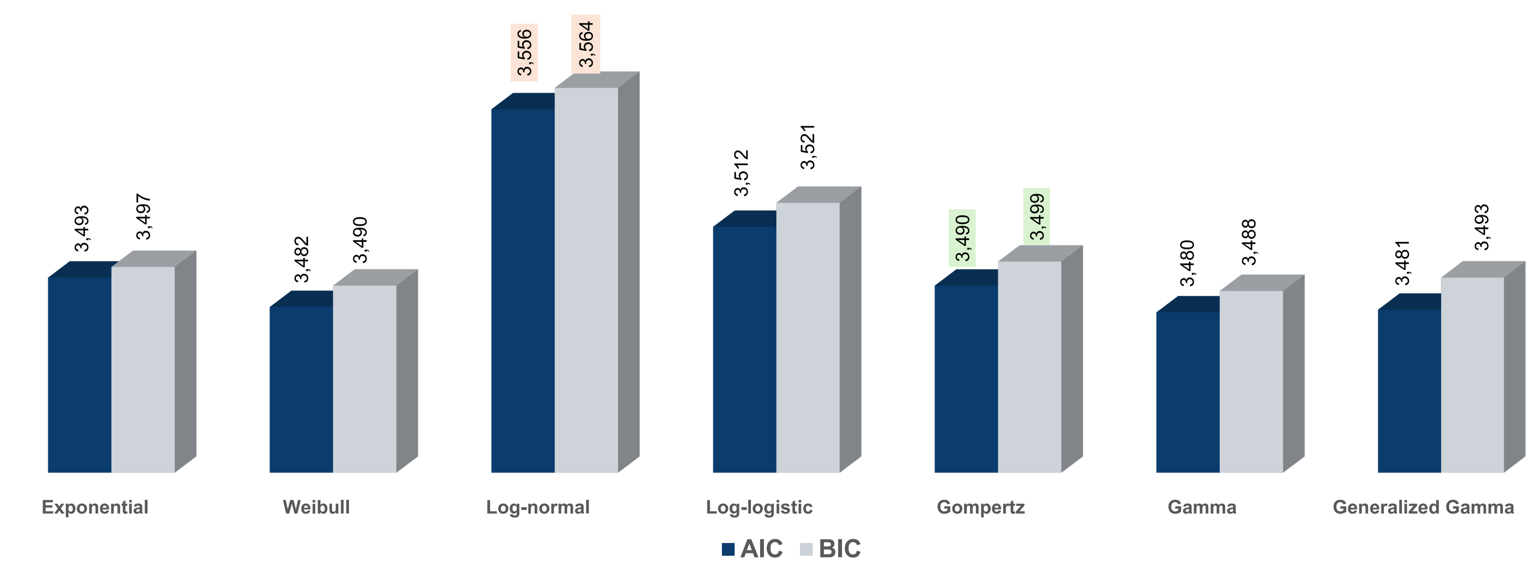
where $f(t)$ represents the overall distribution function, composed of additive component functions $f_k(t)$ for each kth mixture component, and p_k denotes the proportion each component contributes

- The function to estimate the rate/shape/scale parameters was written in R using the *optim* function with box constraints
- The tested weights of two distributions ranged from 1% to 99% with an increment of 1%, resulting in total of **4158** combinations

RESULTS

- Both standard parametric and mixture models was compared based on AIC, BIC and visual inspection to identify the appropriate model for the complex hazard time-to-event data
- Among the parametric combinations tested, the Weibull and Gamma mixture demonstrated a strong fit with AIC and BIC scores of 3,490 and 3,499, respectively.

Figure 1: AIC and BIC for Single Parametric Models



- The Log-logistic and Gompertz mixture model emerged as the optimal choice across all tested models, achieving the lowest AIC and BIC at 3,398 and 3405, respectively
- The Log-logistic got 38% of the model weight while determining some characteristics of subpopulation and the Gompertz got 62% of the model weight

Figure 2: Parametric mixture models fitted on observed data

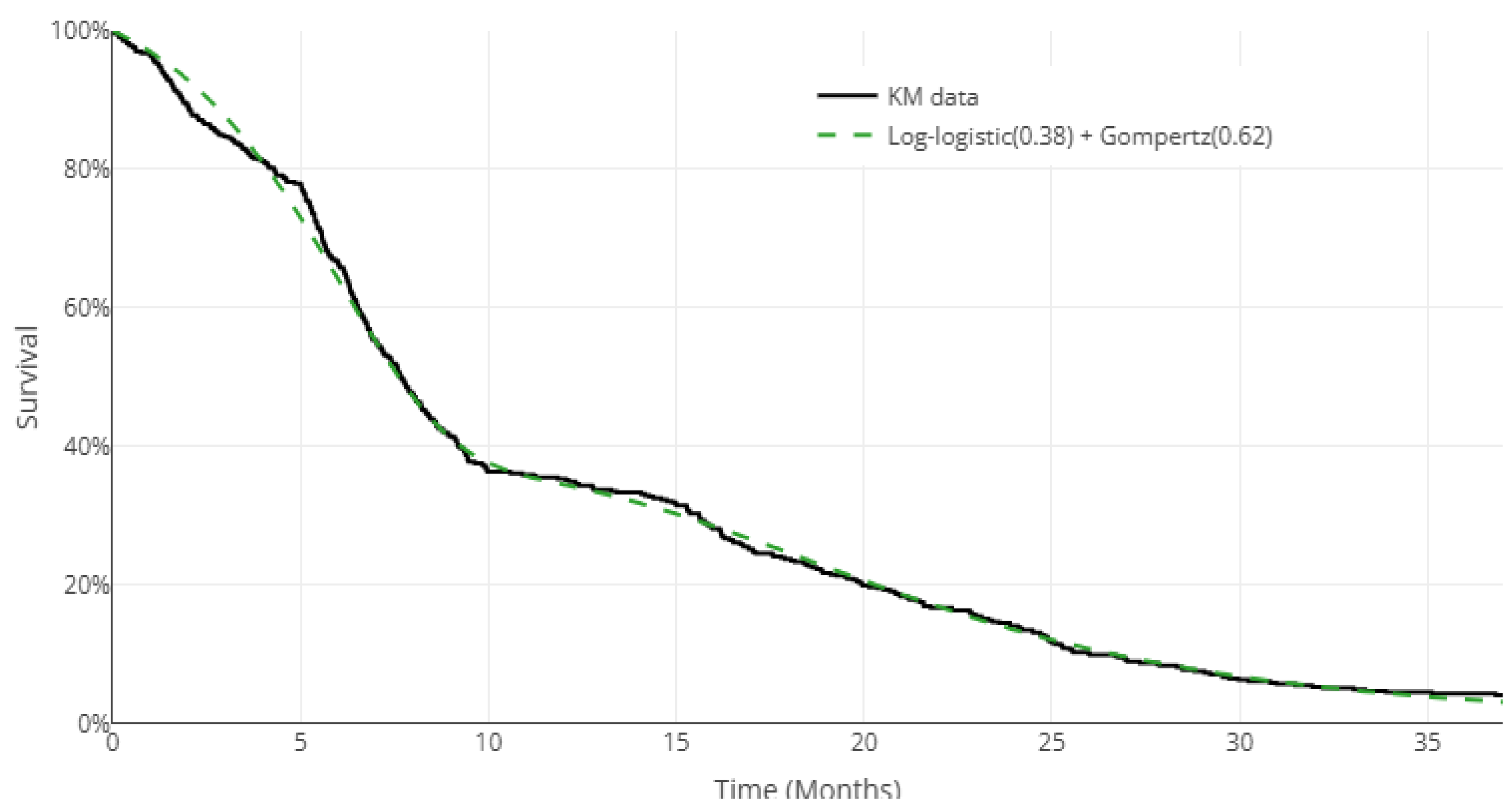


Table 1: AIC and BIC for Parametric Mixture Models with mixture proportions (Top 10)

Model	Mixture 1 proportion	Mixture 2 proportion	AIC	BIC
Log-logistic - Gompertz	38%	62%	3,398.8	3,405.7
Gompertz - Generalized gamma	64%	36%	3,401.1	3,422.4
Gompertz - Log-normal	75%	25%	3,404.2	3,421.0
Log-logistic - Gompertz	39%	61%	3,404.3	3,421.1
Log-normal - Gompertz	39%	61%	3,405.8	3,422.7
Generalized gamma - Gamma	71%	29%	3,407.4	3,428.4
Log-normal - Weibull	23%	77%	3,408.5	3,425.4
Log-normal - Exponential	27%	73%	3,410.4	3,423.0
Weibull - Gompertz	34%	66%	3,413.8	3,430.7
Weibull - Gamma	69%	31%	3,415.0	3,431.8

References

National Institute for Health and Care Excellence (NICE). (2013). *Technical Support Document 21: Flexible survival modeling for cost-effectiveness analyses.* London: NICE Decision Support Unit

Disclosures

SP, PB, BS SK and AS the authors, declare that they have no conflict of interest

Scan to view and save the poster