On When and How to Use External Data to Inform Long-term Survival Curve Extrapolation

X. Gregory Chen[1], Jingshu Wang[2], Kate Young[2]

[1] *MSD Switzerland, Zurich, Switzerland,* [2]*Merck & Co., Inc., Rahway, NJ, USA*

Introduction

One challenge of long-term extrapolations is its external validity, i.e. accuracy of prediction at time points far beyond the observation period when, for example, the prediction survival probability in 30-40 years based on an oncology trial of 3-5 years.

Hence, incorporating external evidence to validate or inform the extrapolation has emerging interests from both payer and sponsor side, particularly when trial data is immature.[1]

Bullement et al 2023 [2], in a systematic review, identified 22 studies for estimating OS that incorporate external data or information. None of them compared or validated their method against another that also incorporated external data or information. This assessment agrees with NICE TSD 21 [3], from January 2020, to reflect the field seem to be still quite green. One framework was repetitively mentioned for its flexibility to cope with different type of external data or info, namely, Guyot et at. [4]

Objectives

Our objective here is to explore when and how to use external data to inform long-term survival extrapolation in a Bayesian framework

The "When"

The "When" refers to when to use external data in model-fitting process for long-term extrapolation rather than as benchmarks for validation.

A brief review of the literature [2,5] combined with our empirical experience identified when it may be necessary to incorporate external data:

- Mixture/non-mixture cure fraction models adjusting for background mortality
- Results are sensitive to the choice of extrapolation model
- Survival data are relatively immature, for example, where median survival has not been reached
- To avoid clearly implausible extrapolations
- When the trial data does not reflect the interested population for reimbursement

Potential sources of external data with hinged pitfalls and key points to consider when assessing the appropriateness of external data sources are presented in Table 1. [2,3,5]

Table 1. Potential Sources of External Data with Key Consideration

Source	Pitfalls	Considerations
General population mortality (national life table estimates for the general population)	▪ Stratified usually only by gender and age, and occasionally by race/ethnicity ▪ Usually taken as without uncertainty in general population adjustment methods	Adjust to match clinical trial characteristics
Other clinical trials	▪ May not include information on newer treatments ▪ May not include sufficiently long follow-up	Generalizability: Similarity or differences in patient populations and other confounding factors
Registry data		
Real-world data		
Clinical opinion	▪ Difficulty of identifying an unbiased point estimate for a new treatment	Generalizability of discrete survival estimates

The "How"

The "How" refers to how to incorporate external data or information into the model fitting process. Either a Frequentist or a Bayesian paradigm could be used. Bayesian framework (e.g. [4,6,7,8]) is more prevalent and it is the focus of this poster.

Key benefits of a Bayesian approach

- All Bayesian inference and prediction is via the posterior, which is a product of the priors and the likelihood. Easier to be explained to the consumers and collaborators, and may be appreciated as a solution with higher transparency
- Easier to control the influence degrees of the external data or information on the model and prediction via the prior setup
- Posterior is typically approximated by simulated data points via a numerical algorithm (e.g. MCMC). Only small additional effort needed to derive uncertainty (e.g. credibility internal) for the interested point estimate or prediction.
- More naturally cope with hierarchical or sequential modeling to allow propagation of insights and uncertainty from an earlier step to the later in a decision process.

Three types of Inputs in model-fitting

Given that it is appropriate and suitable external data or information is available (as laid out in the "When"), these external evidences generally come in or can be transformed into 3 types of input for Bayesian model fitting, see Table 2. It also summarizes how to incorporate each type of input in the Bayesian modeling.

Table 2. Three Types of Input from External Evidences in Model Fitting

	Type of External Data or Information	Typical Source	Data Format	How to incorporate in Bayesian Model-fitting
1	From timepoint t_l to t_u , we expect survival probability $S_{trt}(t)$ in a treatment arm is not better than a given upper bound $\{S^*_{trt}(t_k): t_l \leq t_k \leq t_u, k = 1, 2, \dots, K\}$	Typically, these information is only available for the placebo arm, and based on registry data of the general population that are adjusted for some key characteristics (age, gender, disease history, etc)	A table with two columns, one is the time points $\{t_k\}_{k=1}^{K}$ and the other is the theoretical upper bound $S^*_{trt}(t_k)$ for the survival probability $S_{trt}(t_k)$	$S_{trt}(t_k) = S^*_{trt}(t_k) - \nu$ $\nu \sim F_\nu$ where F_ν is a distribution with only positive value (such as gamma or inv-gamma)
2	From timepoint t_l to t_u , we expect conditional survival probability in a treatment arm $S_{trt}(t\|t-a)$ for some $a > 0$ can be estimated from some external data	Typically, these information is only available for the placebo arm, and based on registry data of the general population that are adjusted for some key characteristics (age, gender, disease history, etc)	A table with at least 4 columns, each row k shows, out of n_k people alive at a time u_k, r_k survive until v_k. (Improved format of external data by [7] comparing to [4])	$r_k \sim Bin(p_k, n_k)$ $p_k = S_{trt}(v_k)/S_{trt}(u_k)$ $p_k \sim Beta(a_0, b_0)$ where a_0 and b_0 could be chosen to be non-informative, e.g. $a_0 = b_0 = 1$
3	Assumption on how the hazard changes over time (continuously or stepwise)	Clinical expert opinion, or a hypothetical scenario with justifiable base	N/A	Bullement et al [2] suggests put a normal prior with very small variance on the hazard ratio after certain time point (says after 6 years). Alternatively, one may set the model such that spline knots after certain time has the same coefficient between arms.

Three Spline Models and Implementation Notes

Standard parametric models (e.g. Weibull, log-logistic, Gompertz) can be quite restrictive in this case. Hence this poster focuses on a class of flexible spline models that can incorporate external data and information.

Three setups of spline model are identified in the literature research. They aim to model different responses and use different spline basis, see the summary with pros and cons in the table below. Since hazard, cumulative hazard, survival probability and likelihood can be derived for one based on the other, these models can **ALL** be used for the common purpose of survival extrapolation as well as incorporate all three types of input listed in the table before.

Table 3. Three Types of Spline Models

Response	Spline Model	Pros and Cons	Implementation *
Log-Cum.-Hazard	Restricted Natural Cubic Spline [4, 9]	**Pros:** Log-Cum.-Hazard is strictly non-decreasing, smoother than hazard or log-hazard curves, hence small number of knots (1-4) suffices to approximate the true curve. In addition, little effort to transfer log-cum.-hazard into conditional survival probability and likelihood, hence computationally efficient in a Bayesian approach. **Cons:** Without restriction, posterior sample via a numerical algorithm would contain illegible sample of spline parameters that leads to decreasing or wiggly curves that violates the theoretical expectation. This introduces arbitrary bias in the final point estimate, if posterior mean or median are used, and their uncertainty estimates.	Example code from [4, 6] using WinBUGs. Code from [4] needs refinement before reproducing the result in the paper using {rjags}, and non-convergence were observed for several parameters. No R package is available. This means high technical expertise and effort are needed for implementation.
Hazard	M-Spline [7, 8]	**Pros:** Hazard is strictly positive, so as the response of a M-spline model, hence no illegible posterior sample is expected. **Cons:** Number and location of knots should be more carefully treated than Hazard needs numerical integration to form likelihood, but due to the setup of M-Spline, the integration can be done analytically and hence can alleviate the computational burden.	{survextrap} Version 0.8.6 (2023/10/24), a Beta version, Can handle type 2 of external data or information in Table 2. Not clear how to incorporate the other two types.
Log-Hazard	B-Spline [8]	**Pros:** log-hazard can be both positive or negative, so as the response of a B-spline model, hence no illegible posterior sample is expected. **Cons:** Similar to M-spline model above. In addition, numerical integration to form likelihood in this case would have no short cut. So this model is likely to be the most computationally expensive.	{rstanarm} Version 2.26.1 (2023/9/13), an official version Can handle type 2 of external data or information in Table 2. Not clear how to incorporate the other two types.

One More Key Issue: Location of Knots

Number of knots and their locations are crucial parameters in a spline model. It impacts the goodness-of-fit and often also the convergence of MCMC algorithm. In the Bayesian setting, these parameters are rarely up for inference (i.e. no prior for them). Among difficulties to derive correct posterior samples given various type of external data, this additional layer of complexity or Bayesian model selection are typically put secondary, and hence overlooked in the current literature.

Conclusion and Discussion

There are many factors to be considered before incorporating external data directly into survival model for long term extrapolation. Once deemed appropriate, a more systematic approach is needed to incorporate different types of external data and info in a Bayesian framework. More methodological research in this area is greatly needed.

Reference:

1.Latimer N. NICE DSU Technical Support Document 14: Survival Analysis for economic Evaluations Alongside Clinical Trials – Extrapolation with Patient-level data. 2013. Decision Support Unit, University of Sheffield, UK.

2.Bullement, A., Stevenson, M.D., Baio, G., Shields, G.E. and Latimer, N.R., 2023. A systematic review of methods to incorporate external evidence into trial-based survival extrapolations for health technology assessment. *Medical Decision Making*, p.0272989X231168618.

3.Rutherford MJ, Lambert PC, Sweeting MJ, et al. NICE DSU Technical Support Document 21: Flexible Methods for Survival Analysis. 2020. Decision Support Unit, University of Sheffield, UK.

4.Guyot, P., Ades, A.E., Beasley, M., Lueza, B., Pignon, J.P. and Welton, N.J., 2017. Extrapolation of survival curves from cancer trials using external information. *Medical Decision Making*, 37(4), pp.353-366.

5.Palmer S, Borget I, Friede T, et al. 2023. A Guide to Selecting Flexible Survival Models to Inform Economic Evaluations of Cancer Immunotherapies. *Value in Health* 26:185-192

6.Vickers, A., 2019. An evaluation of survival curve extrapolation techniques using long-term observational cancer data. *Medical Decision Making*, 39(8), pp.926-938.

7.Jackson, C., 2023. survextrap: a package for flexible and transparent survival extrapolation. arXiv preprint arXiv:2306.03957.

8.Brilleman, S.L., Elci, E.M., Novik, J.B. and Wolfe, R., 2020. Bayesian survival analysis using the rstanarm R package. arXiv preprint arXiv:2002.09633.

9.Royston, P. and Parmar, M.K., 2002. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in medicine*, *21*(15), pp.2175-2197.