Checkmate 067 study Bayesian mixture cure model analysis

Nathan Green, Gianluca Baio, UCL

7 January 2021

Executive summary

In this project we formulate and demonstrate the application of a Bayesian mixture cure model (MCM) using the Checkmate 067 study dataset and the Exponential and log-logistic distributions for event times. Analogous results to those created previously for the frequentist MCM approach are produced and we extend the Bayesian MCM to incorporate additional structure, including the joining of overall survival (OS) and progression-free survival (PFS) models using a hierarchical struture on the cure fraction. We show that the Exponential and in particular the log-logistic OS and PFS Bayesian MCMs perform reasonably well for the Checkmate 067 data. The real benefit of this approach may be with other dataset where there is short follow-up or small sample sizes. The associated R code for this work, held in a private on-line repository, has been written for re-use and generalisability to other problems.

Background

Immuno-oncologic (IO) studies for melanoma therapies, such as *ipilimumab* (`ipi`), *nivolumab* (`nivo`), and the *nivolumab* with *ipilimumab* (`nivo + ipi`) combination, have indicated that survival curves "plateau" (a considerable proportion of patients are "long-term survivors"). Cure models are a special type of survival analysis where this "cure fraction" (the underlying proportion of responders to treatment/long-term survivors) is accounted for. Cure models estimate the cure fraction, in addition to a parametric survival function for patients that are not cured. The mortality risk in the cured patients is informed by a background mortality rate. The population that is not cured is subject both to background mortality and to additional mortality from their cancer, estimated using a parametric survival model.

A mixture cure model (MCM) (Amico and Van Keilegom (2018)) is a type of cure model where survival is modelled as a mixture of two groups of patients: those who are cured and those who are not (and who therefore remain at risk). The survival for a population with a cure fraction can be written as follows:

$$S(t,x) = S^*(t,x)[\pi(x) + (1-\pi(x))S_u(t,x)], \qquad (*)$$

where $S(t,x)$ denotes the survival at time t, $S^*(t,x)$ denotes the background mortality at time t conditional on covariates x, $\pi(x)$ denotes the probability of being cured conditional on covariates x, and $S_u(t,x)$ denotes the event (progression or mortality) due to cancer at time t conditional on covariates x. For PFS, the survival is composed of either progressing to a disease state or death.

Aims

The aims of the the analysis in this document are as follows:

- Demonstrate the application of a Bayesian mixture cure model using the Checkmate 067 study dataset and the Exponential distribution for event times.
- Produce analogous results to those created previously for the frequentist approach.
- Extend the Bayesian model to incorporate additional structure including a hierarchical cure fraction.

This analysis has been carried-out using the Stan inference engine (Carpenter et al. (2017)) called from R on a Windows PC. The packaged code can be downloaded from a private GitHub repository with permission from the package authors at https://github.com/StatisticsHealthEconomics/rstanbmcm. See the *How to use rstanbmcm* vignette for an introduction to how to use the package.

Likelihood

Let T_i be the non-negative random variable denoting the survival time of patient i with covariate vector x_i.

In the simplest case we can assume that the cure fraction is the same for the whole population i.e. π is fixed. Further, we can assume the π models the relationship between x_i and the probability of being cured. E.g. using a logistic-linear model

$$\pi(x_i|\beta) = 1/[1 + \exp(-x_i^T\beta)].$$

The likelihood of the standard survival is

$$L = \prod_i S(t_i|x_i)h(t_i|x_i)^{\delta_i}$$

Log-likelihood is therefore

$$l = \sum_i \log(S(t_i|x_i)) + \delta_i \log(h(t_i|x_i))$$

Plugging this directly into the mixture cure equation in (*) gives

$$l(\pi|\delta, x) = \sum_i \log(S^*(t_i|x_i)h^*(t_i|x_i)^{\delta_i}[\pi(x) + (1 - \pi(x))S_u(t_i|x_i)h_u(t_i|x_i)^{\delta_i}])$$

If we assume that the cured component is the Exponential survival model then the non-cured component can be thought of in similar terms to the cumulative incidence function. That is, the probability of an event is the combined probability of surviving both events (e.g. for OS, all-cause and cancer mortality) and then experiencing either i.e. dropping the S dependencies for brevity

$$S^*S_u(h^*)^\delta + S^*S_u(h_u)^\delta = S^*S_u(h^* + h_u)^\delta \qquad (**)$$

Bayesian formulation

In a Bayesian approach to modelling, all quantities that are subject to uncertainty are modelled using probability distributions. This applies to observed data (e.g. time to PFS for a given individual), that are subject to sampling variability, as well as to unobservable parameters (e.g. the coefficient quantifying the impact of age or sex over the average survival curve). In this latter case, probability distributions are used to model the epistemic uncertainty (e.g. the fact that we do not know for certain what the "true" underlying value of the model parameter is). In addition, we may model as yet unobserved (but potentially observable) quantities using a suitable probability model. For example, we could consider the extrapolated part of the survival curve as subject to uncertainty due to the current sampling process giving rise to the data that are actually observed, as well as the uncertainty on the underlying data generating process.

We can mix different sources of evidence to form our "prior" distributions, which are used to describe the state of science on the model parameters. These are then combined with any observed data to form an updated level of knowledge. This process is particularly relevant in the case at hand, when data can only inform about limited aspects of the overall underlying reality. For this reason, it is important to a) include information/evidence available in the form of external data and/or expert opinion; b) extract the most information possible from the available data (e.g. by formally trying to model the correlation between the PFS and the OS data to borrow strength from the more mature set of observations).

A built-in advantage of the Bayesian procedure is that uncertainty is directly and formally propagated to an economic model; the main output from the statistical analysis (the extrapolated survival curve) are produced by default as based on a full posterior distribution. From this, we can easily derive a "base case" (e.g. taking the mean value) but without the need for further tools (such as bootstrap) we already have a full characterisation of the underlying uncertainty that can be used in the process of probabilistic sensitivity analysis. We can moreover add information in the priors to ensure that the extrapolation beyond the observed data is realistic and consistent with the clinical expertise (e.g. by "anchoring" the extrapolated survival curve to be probabilistically below the curves for the healthy population, or by ensuring that OS behaves in a way to respect some agreed level of similarity, or correlation, to PFS).

Posterior equation Using the likelihood function defined above and prior distributions on uncertain parameters, we can specify the posterior distribution. Defining g_2 as the prior distribution for the coefficients of the uncured fraction β^u and g_3 as the prior distribution for the coefficients of the cured fraction β^*, then the general form of the posterior distribution can be written as follows.

$$p(\pi, \beta^u, \beta^* | \delta, x) \propto L(\pi, \beta^u, \beta^* | \delta, x) f(\pi) g_2(\beta^u) g_3(\beta^*)$$

assuming that the cure fraction is independent of the covariates.

Cure fraction

There are two obvious ways to represent the uncertainty about the cure fraction in the model.

The first is to specify the cure fraction directly using a $\pi \sim Beta(a_{cf}, b_{cf})$ prior, most uninformative as a uniform $Beta(1, 1)$. The parameters can be obtained via transformation of mean and standard deviation to allow a more natural scale for elicitation.

Alternatively, we may specify the uncertainty on the real line with a Normal distribution and then transform to the probability scale.

A further consideration is how to represent the cure fraction so to share information between the OS and PFS data. We will investigate 3 alternatives.

- *Pooled*: Assume that the cure fraction is the same for OS and PFS i.e. $\pi_{os} = \pi_{os} = \pi$ where

$$logit(\pi) \sim N(\mu_{cf}, \sigma^2_{cf}),$$

- *Separate*: Model each independently.

$$logit(\pi_{os}) \sim N(\mu_{cfos}, \sigma^2_{cfos}), \;\; logit(\pi_{pfs}) \sim N(\mu_{cfpfs}, \sigma^2_{cfpfs})$$

- *Hierarchical*: Assume exchangeability between OS and PFS

$$\pi \sim N(\mu_{cf}, \sigma^2_{cf}), \;\; logit(\pi_{os}) \sim N(\pi, \sigma^2_{cfos}), \;\; logit(\pi_{pfs}) \sim N(\pi, \sigma^2_{cfpfs})$$

Below is an example DAG for the hierarchical cure fraction without a joint time to event component. Notice that even without the direct relationship between PFS and OS there is still an indirect influence via π.

Background survival

The previous frequentist analysis used the World Health Organization (WHO) life tables by country for the latest year available of 2016 (WHO (2020)) to inform the background mortality rate (baseline hazard). These baseline hazards are the expected mortality rate for each patient at the age at which they experience the event. The mortality data are age- and gender adjusted, thus providing a granular account of the different patient profiles in the trial. The WHO reports conditional probabilities of death in 5-year intervals until age 85. A constant annual mortality rate is reported for individuals over 85. They assumed that the maximum age is 100 years.

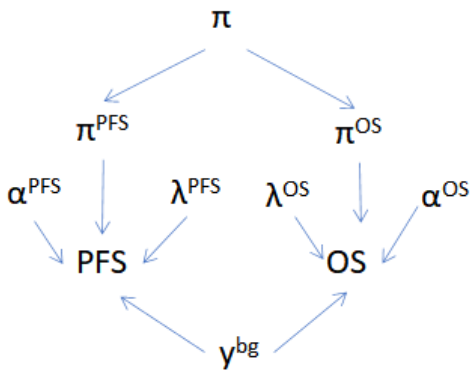

Figure 1: Hierarchical cure fraction DAG.

In a Bayesian analysis there are alternative ways in which we could model the background mortality.

For this work we shall use WHO hazard point estimates as known. We could consider the WHO estimates to provide sufficiently accurate estimates given the sample size and so incorporating uncertainty is not necessary. This also forces consistency across fits. Denote the WHO estimates for individual i as $\hat{f}_i, \hat{S}_i, \hat{h}_i$ for the density, survival and hazard respectively.

Exponentially distributed event times

Consider the straightforward case where both the cancer times to event follow Exponential distributions.

Define $f(t)$ density, $S(t)$ survival and $h(t)$ hazard functions.

$$f(t) = \lambda \exp(-\lambda t), \quad S(t) = \exp(-\lambda t), \quad h(t) = \lambda$$

Which gives the likelihood

$$L(\pi, \lambda_u | \delta, x, \hat{S}, \hat{h}) = \sum_i \hat{S}_i \hat{h}_i^{\delta_i} [\pi(x) + (1 - \pi(x)) \exp(-\lambda_u t) \lambda_u^{\delta_i}])$$

Substituting $S(t)$ and $h(t)$ into (**)

$$f_u^* = e^{-\lambda^* t} e^{-\lambda_u t} (\lambda^* + \lambda_u)^{\delta} \quad i.e. \text{for no censoring } \; T \sim Exp(\lambda^* + \lambda_u)$$

Log-logistic distributed event times

In the frequentist analysis, generally the log-logistic is the one of the best fitting distributions. The relevant functions are

$$f(t) = \frac{(\beta/\alpha)(t/\alpha)^{\beta-1}}{(1 + (t/\alpha)^{\beta})^2}, \quad S(t) = (1 + (t/\alpha)^{\beta})^{-1}, \quad h(t) = \frac{(\beta/\alpha)(t/\alpha)^{\beta-1}}{1 + (t/\alpha)^{\beta}}$$

which gives the likelihood

$$L(\pi, \alpha, \beta | \delta, x, \hat{S}, \hat{h}) = \sum_i \hat{S}_i \hat{h}_i^{\delta_i} \left[\pi(x) - (1 - \pi(x))(1 + (t/\alpha)^{\beta})^{-1} \left(\frac{(\beta/\alpha)(t/\alpha)^{\beta-1}}{1 + (t/\alpha)^{\beta}} \right)^{\delta_i} \right]$$

Results

We fit the exponential hazard model to the study data and produced the posterior survival curves below.

For each model and treatment we produce two figures:

1. The expected survival curves with 95% Credible Intervals (CrI). The OS curves are to the left-hand side and PFS curves to the right-hand side. Background mortality (i.e. cured patients) is indicated by the red line. Non-cured patients survival curves are shown in dark green and blue for OS and PFS respectively. Light green and magenta are the total sample. The black line is the Kaplan-Meier curve for the observed data. Note that these plots are for an average individual, e.g. at average age, and so we would not expect them to perfectly match the sample data Kaplan-Meier.
2. Kaplan-Meier curves for 50 simulated trials using the posterior prediction distribution for time to event. This is a good model checking plot, indicating if we can replicate data similar to the observed data using the fitted model.

Figures 2, 3, and 4 show posterior survival curves for the Exponential and log-logistic models.

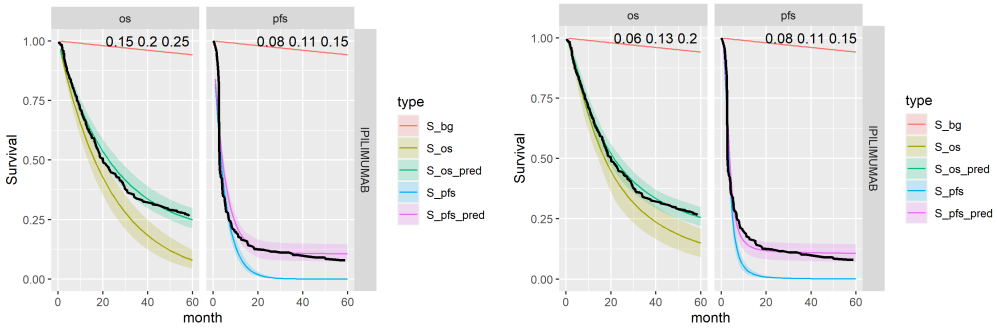

Figure 2: Posterior survival curves for the mixture cure model and ipilimumab. The red line is cured background, light green and blue are uncured, and dark green and magenta are combined. a) Exponential uncured; b) Log-logistic uncured.

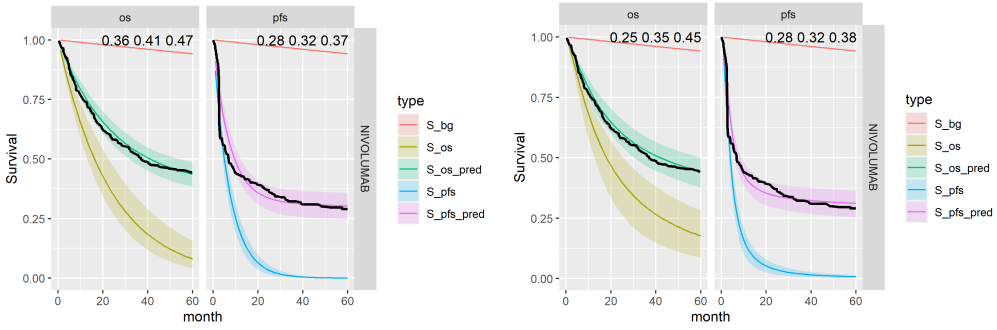

Figure 3: Posterior survival curves for the mixture cure model and nivolumab. The red line is cured background, light green and blue are uncured, and dark green and magenta are combined. a) Exponential uncured; b) Log-logistic uncured.

Figure 5 show the forest plot of cure fraction posterior distributions. We see that the values are generally similar for the Exponential and log-logistic fits. This clearly shows how the global cure fraction posterior distribution lies partway between the PFS and OS distributions.

The table below summarises the cure fraction posterior distribution for each scenario.

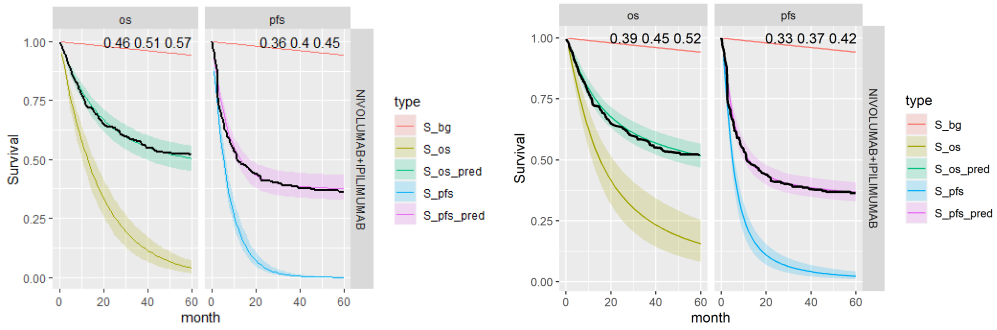

Figure 4: Posterior survival curves for the mixture cure model and dual ipilimumab and nivolumab. The red line is cured background, light green and blue are uncured, and dark green and magenta are combined. a) Exponential uncured; b) Log-logistic uncured.

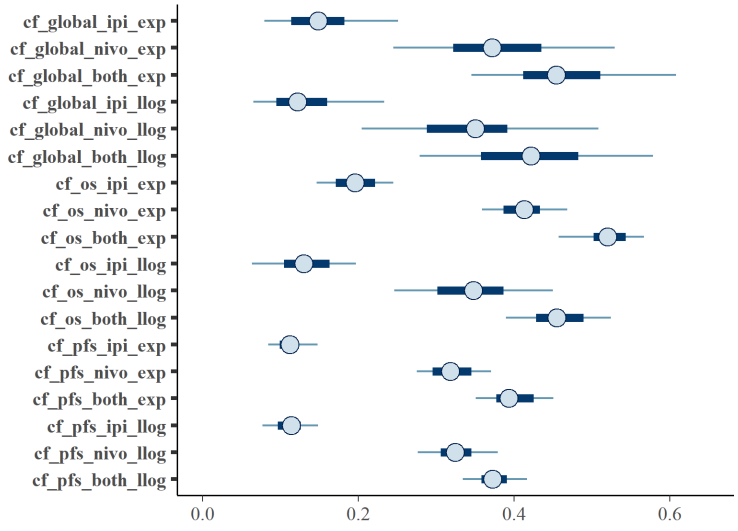

Figure 5: Forest plot of cure fraction posterior distributions.

	OS D istn	PFS D istn	Treatment	cf (CrI)	cf_{OS} (CrI)	cf_{PFS} (CrI)
1	exp	exp	IPILIMUMAB	0.151 (0.068, 0.26)	0.196 (0.136, 0.255)	0.112 (0.079, 0.155)
2	exp	exp	NIVOLUMAB	0.378 (0.216, 0.557)	0.412 (0.35, 0.479)	0.321 (0.262, 0.379)
3	exp	exp	NIVOLUMAB + IPILIMUMAB	0.463 (0.322, 0.62)	0.519 (0.448, 0.585)	0.398 (0.336, 0.455)
4	log logi stic	log logi stic	IPILIMUMAB	0.133 (0.061, 0.278)	0.133 (0.054, 0.206)	0.113 (0.075, 0.152)
5	log logi stic	log logi stic	NIVOLUMAB	0.348 (0.191, 0.526)	0.345 (0.231, 0.456)	0.325 (0.264, 0.382)
6	log logi stic	log logi stic	NIVOLUMAB + IPILIMUMAB	0.421 (0.266, 0.599)	0.459 (0.37, 0.535)	0.375 (0.327, 0.421)

The table below gives the leave-one-out cross validation statistics for each model fit.

OS Distn	PFS Distn	Treatment	Statistic	Estimate	SE
exp	exp	IPILIMUMAB	elpd_waic	-1838.83	36.07
exp	exp	NIVOLUMAB	elpd_waic	-1679.59	45.69
exp	exp	NIVOLUMAB+IPILIMUMAB	elpd_waic	-1543.24	53.80
loglogistic	loglogistic	IPILIMUMAB	elpd_waic	-1768.17	39.20
loglogistic	loglogistic	NIVOLUMAB	elpd_waic	-1651.73	46.02
loglogistic	loglogistic	NIVOLUMAB+IPILIMUMAB	elpd_waic	-1533.95	53.17
exp	exp	IPILIMUMAB	p_waic	5.37	0.48
exp	exp	NIVOLUMAB	p_waic	7.45	0.62
exp	exp	NIVOLUMAB+IPILIMUMAB	p_waic	6.67	0.52
loglogistic	loglogistic	IPILIMUMAB	p_waic	9.19	0.56
loglogistic	loglogistic	NIVOLUMAB	p_waic	9.20	0.53
loglogistic	loglogistic	NIVOLUMAB+IPILIMUMAB	p_waic	6.10	0.50
exp	exp	IPILIMUMAB	waic	3677.65	72.13
exp	exp	NIVOLUMAB	waic	3359.19	91.37
exp	exp	NIVOLUMAB+IPILIMUMAB	waic	3086.47	107.61
loglogistic	loglogistic	IPILIMUMAB	waic	3536.33	78.40
loglogistic	loglogistic	NIVOLUMAB	waic	3303.47	92.04
loglogistic	loglogistic	NIVOLUMAB+IPILIMUMAB	waic	3067.91	106.35

The variance partition coefficient (VPC) is defined as $\sigma^2_{global}/(\sigma^2_{global} + \sigma^2_e)$ where $e = PFS$ or OS. This indicates want proportion of the total variance is attributable to variation within-groups, or how much is found between-groups.

OS Distn	PFS Distn	Treatment	PFS	OS
exp	exp	IPILIMUMAB	0.813	0.784
exp	exp	NIVOLUMAB	0.876	0.877
exp	exp	NIVOLUMAB+IPILIMUMAB	0.857	0.846
loglogistic	loglogistic	IPILIMUMAB	0.812	0.604
loglogistic	loglogistic	NIVOLUMAB	0.890	0.672
loglogistic	loglogistic	NIVOLUMAB+IPILIMUMAB	0.915	0.786

It appears the above table that much of the variation is due to between PFS and OS indicating distinct cure fractions in each.

References

Amico, Maïlis, and Ingrid Van Keilegom. 2018. "Cure Models in Survival Analysis." *Annual Review of Statistics and Its Application* 5: 311–42. https://doi.org/10.1146/annurev-statistics-031017-100101.

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76 (1).

WHO. 2020. "Life tables by country." http://apps.who.int/gho/data/?theme=main&vid=61160.