Bayesian Mixture Cure Modelling in Metastatic Melanoma (CheckMate 067)

N Green, UCL

14/10/2020

[taken from Anthonio]

### Methodology

Immuno-oncologic (IO) studies for melanoma therapies, such as *ipilimumab*, *nivolumab*, and the *nivolumab* + *ipilimumab* 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) 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:

where denotes the survival at time , denotes the background mortality at time conditional on covariates , denotes the probability of being cured conditional on covariates , and denotes the mortality due to cancer at time t conditional on covariates . We use World Health Organization (WHO) life tables by country (2018) to inform the background mortality rate (baseline hazard) inputed to the cure model. Such baseline hazard is the expected mortality rate for each patient at the age at which he/she experiences the event. The mortality data are age- and gender adjusted, thus providing a granular account of the different patient profiles in the trial. Note that WHO reports conditional probabilities of death in 5-year intervals until age 85. A constant annual mortality rate is reported for individuals over 85. In addition, we make the assumption that the maximum age that a patient can live up to is 100 years.

To model the disease-specific mortality (the uncured fraction), all standard parametric distributions are tested:

* exponential
* Weibull
* Gompertz
* log-normal
* log-logistic
* generalised gamma.

Parameters for the models, including the cure rate parameter, are derived via Bayesian inference.

### Model description

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

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

The likelihood of the standard survival is

Log-likelihood

Using this with the mixture cure equation gives

Posterior distribution with cure fraction independent on covariates.

Cure fraction dependent on covariates.

### Exponential

Define density, survival and hazard functions.

Which gives the likelihood

#### Prior specification

The marginal prior distributions will probably depend on which group of data are being modelled.

* *Cure fraction*

We can specify directly using a Beta(, ) prior, most uninformative at a uniform Beta(1,1). Incorporating the covariate via a logistic link we specify priors on the linear coefficient . Although not strictly recommended, we could take sensible starting values from the previous frequentist analysis results. For example, for OS the nivo, ipi and combined mean cure fractions were 45%, 23% and 54%. If the data are centred then this corresponds to , respectively. We would still need to specify how much we expect one year increase in age would change the cure fraction.

* *Background survival*

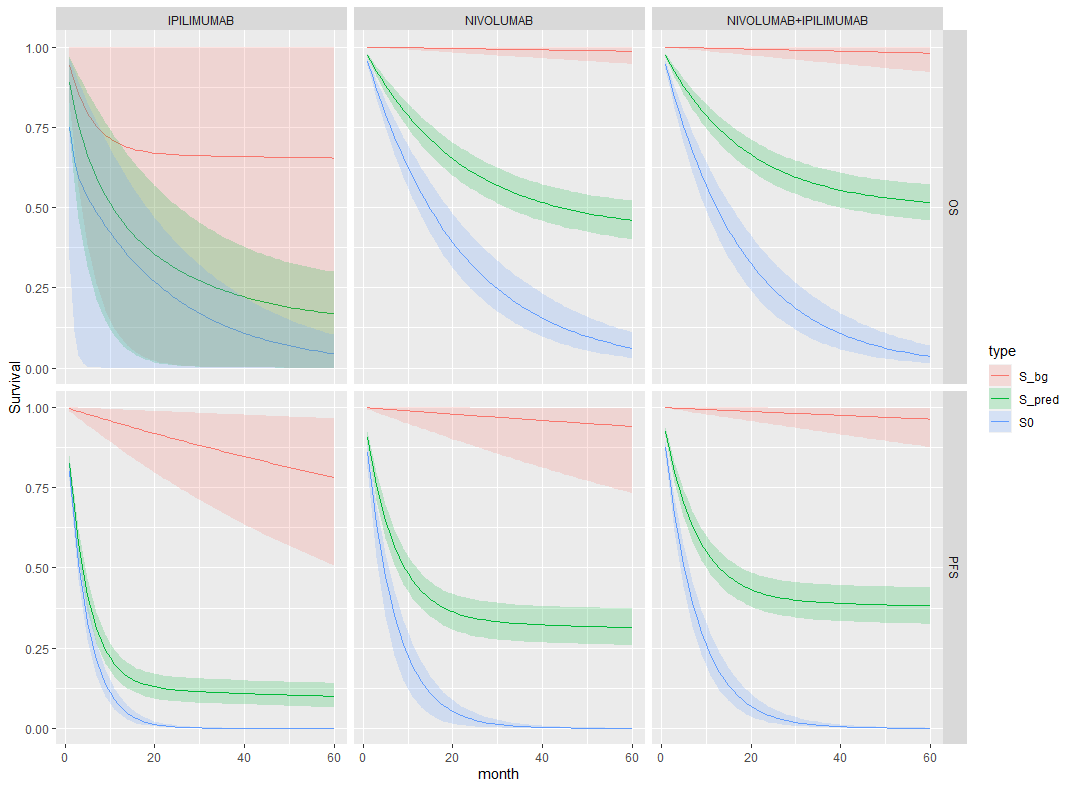
The default survival curves used in this analysis are from country-level WHO data. We can consider these to provide sufficiently accurate estimates given the sample size and so incorporating uncertainty is not necessary. In fact these annual point estimates are used directly in the frequentist analysis. This is done by age matching each individual with the life table and then obtaining the mortality rate at the event time.

However, we also want the developed model to be able to be applied to other datasets which may be smaller or noisy. Sensible prior parameter values can be taken for the life table hazard curve. After infancy the log-hazard is approximately linear and so intercept and slope estimates are simple to obtain.

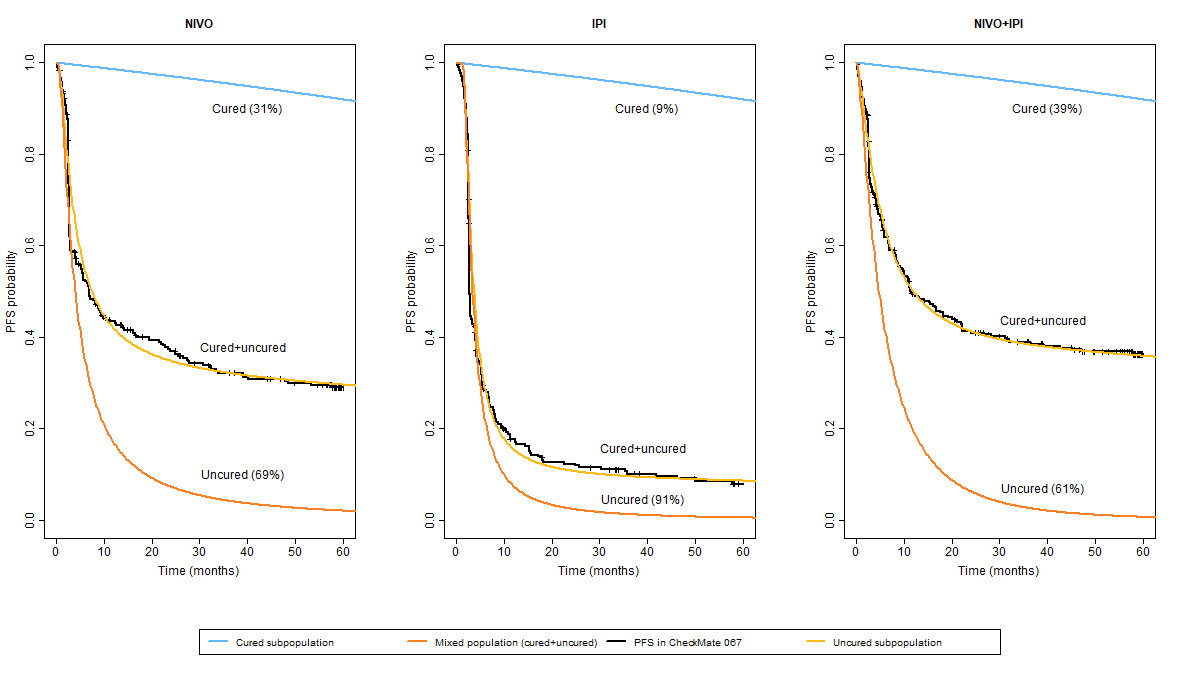
* *Mortality due to cancer*

### Initial results

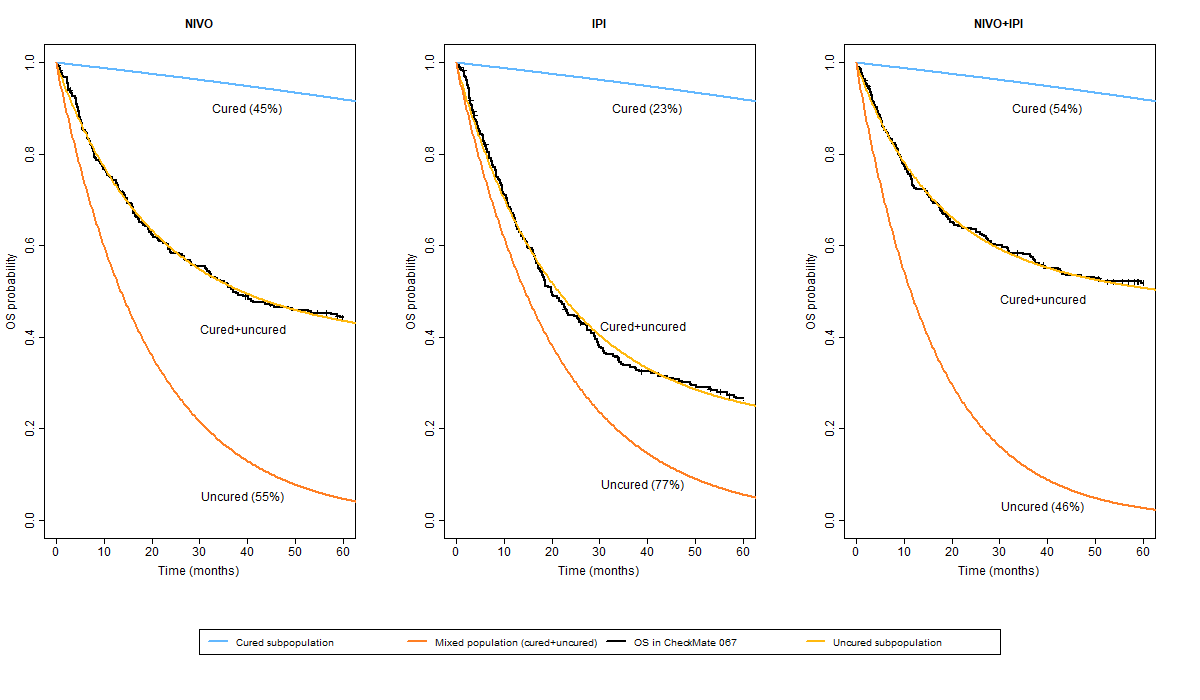
We fit the exponential hazard model to the study data and produced the posterior survival curves below. The corresponding frequentist curves are also shown for comparison. We see that the curve are similar, providing some initial validation. Note that one the fits has clearly not converged and so the plots is not meaningful. This is an indication that we need to improve how the model is set-up prior to fitting. This could be that the exponential distribution is not appropriate or that better prior specification of the parameters is needed. We ran the stan engine for over 70,000 iterations so a brute force approach of simply increasing the number of iterations may not be appropriate in this case.



Exponential hazard posterior survival



Frequentist approach



Frequentist approach