Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Bayesian hierarchical mixture cure modelling †

Nathan Green*[1] | Gianluca Baio[2,3] | Author Three[3]

[1]Department of Statistical Science, UCL, London, UK
[2]Org Division, Org Name, State name, Country name
[3]Org Division, Org Name, State name, Country name

Correspondence
*Nathan Green Email: n.green@ucl.ac.uk

Present Address

Summary

Time to an event of interest over a lifetime is a central measure of the clinical benefit of an intervention used in a health economic evaluation (HEE). Within the same trial multiple event-types may also be considered. For example, overall and progression-free survival time for different drugs in oncology studies. A common situation is when an intervention is only effective for some unobserved proportion of the population. So we need to estimate latent group membership as well as separate survival models for each. However, in real-life trials follow-up may be relatively short leading to substantial censoring. We present a general Bayesian hierarchical framework that can handle this complexity by exploiting the similarity of cure fractions between event types; accounting for the correlation between them and improving the extrapolation beyond the observed data. Assuming exchangeability between cure fractions facilitates the borrowing of information between event type. We show the benefits of using our approach with a motivating example, the Checkmate 067 trial.

KEYWORDS:
Bayesian, survival analysis, oncology

1 | INTRODUCTION

Better diagnosis and screening has lead to more cancer patients with long-term survival (LTS) equivalent to the general population. In previous immuno-oncologic (IO) studies for melanoma therapies, such as *ipilimumab*, *nivolumab*, and dual *nivolumab* and *ipilimumab* combination, results have indicated that survival curves "plateau" for a considerable proportion of LTS patients [ref]. This proportion of patients is often called the *cure fraction*. The term 'cured' here means that a patient has responded to treatment and is a LTS.

Cure models are a special type of survival analysis where this cure fraction is accounted for. Cure models split patients in to two groups (cured or not) and individuals are subject to one or two sources of risk. 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 disease, estimated using a survival model.[?] There are several types of cure models and a rich and growing body of research in this area. (see[?] for a comparison and guidance with application to oncology.)

Cure models are important because ... if you don't account for a mixture then give the wrong results

Specifically when applied in health economics modelling, cure models are important because the results of the life time survival are used in cost-effectiveness analyses which influence policy decisions. Incorrect modelling will misinform the decision-makers.

BMC to give background...

†title footnote.
[0]**Abbreviations:** MCM, mixture cure model

Several issues with existing approaches:

Firstly, when the event-types are correlated, such as with OS and PFS, this should be accounted for in the modelling. Performing separate, independent analysis may give counter-intuitive results. In the case of multiple types of end-points in the same trial, since the clinical interpretation of a cure fraction is the proportion of patients that respond to treatment, it may make intuitive sense that there is a single underlying cure fraction. For example, what would be the interpretation if there are differences in emergent survival plateaus between PFS and OS? This is a clinically unintuitive dichotomy between the resulting proportions of long-term survivors (LTS).

Secondly, time to event data is classically right-censored times due to end to follow-up. This may be too premature to adequately characterise the survival model for a single event-type but by thinking about the set of event-types together this may be alleviated.

Lastly, extrapolation out to life times of patients is important for health economics modelling in order to guage the full effect of an intervention.

Building on the existing literature, the main contribution of this paper is to address all of these issues. A Bayesian hierarchical model of PFS and OS, which has the double advantage of borrowing information between event types (e.g. the likely more mature PFS data to inform the highly censored OS) and obtaining a global cure fraction.

For a non-cure fraction model survival analysis in health economics[??]. A frequentist multi-level modelling approach in mixture cure modelling has been studies previously[?]. They use random effects to model multilevel clustering structure in the linear predictors in both hazard function and cured probability parts. However, to our knowledge there has been no Bayesian fully-parametric mixture cure model with a multi-level modelling structure for event types. Correlation between random effect in the uncured survival and cure fraction has also been investigated using a bivariate Normal distribution[?].

A suite of distributions given in NICE guidelines for Health Technology Assessments (HTA)[?] include Exponential, Weibull, Gompertz, Log-logistic and Log-Normal. This has been interpreted as meaning all of these distributions should be modelled with a given data set. This prescriptive approach does not take into account what we know *a priori* about the problem. In reality, a subset of these distributions will be more appropriate for the problem.

In many mixture cure fraction analyses the survival associated with the uncured fraction of the population is represented by a Cox proportional hazard model. This means that the baseline hazard is left unspecified but still the coefficients can be estimated. However, we are concerned with the application of cure models to a health economics context. In this case it is important to determine life-time survival in order to calculate total health impact measures such as quality-adjusted life-years (QALY) and life-years lost (LYL). Therefore, the semi-parametric approach is inadequate because we must extrapolate beyond the observed data. In this paper we will adopt a fully-parametric model for the uncured survival.

This analysis was carried-out using the Stan inference engine[?] called from R[?] using the package `rstan` on a Windows 10 PC. Each fit employed 4 chains of 1000 iterations each, and a burn-in of 100 iterations. The details of the algorithm can be found in the Appendix. The packaged code is in a generalisable framework and can be downloaded from here.

The paper is structured as follows. Next, the motivating example and data used throughout is outlined in Section 2. We introduce the basic mixture cure model in Section 3.1, before extending this to our Bayesian hierarchical model in Section 3.2. In Section 4 our novel modelling approach is applied to the motivating example dataset. Comparisons are made between our approach and an independent model analogue. Finally, in Section 5 we discuss the results and future work.

2 | MOTIVATING EXAMPLE

Our motivating example concerns a long term study of melanoma therapies [ref].

2.1 | Checkmate 067 trial data

Figure B1 shows the Kaplan-Meier survival curves for the OS and PFS study data.

[**FIGURE** 1 about here.]

The data source for the analysis is the patient-level data from the CheckMate 067 trial [ref?]. Our dataset contains $n = 945$ subjects (8 with a missing treatment indicator, i.e. the actual treatment received). The treatment groups are as follows:

Nivolumab monotherapy 3 mg/kg intravenous (IV) once every 2 weeks (Q2W). 313 patients were treated (203 PFS events and 175 OS events). Median PFS and OS are 6.93 months (95% confidence interval (CI) 5.32 - 10.41) and 36.9 months (95% CI 31.24 - 60.9), respectively. *should time be in months/days?*

Ipilimumab monotherapy 3 mg/kg IV once every 3 weeks (Q3W) for a total of 4 doses. 311 patients have been treated (261 PFS events and 228 OS events). Median PFS and OS are 2.86 months (95% CI 2.79 - 3.29) and 20.0 months (95% CI 17.22 - 25.6), respectively.

Combined nivolumab with ipilimumab 1 mg/kg IV and 3 mg/kg IV Q3W for 4 doses followed by nivolumab 3 mg/kg IV Q2W. 313 patients were treated (182 PFS events and 151 OS events). Median PFS is 11.50 months (95% CI 9.26 - 20.80) and median OS not reached.

A full set of summary statistics can be found in the supplementary material.
How it compares with dataset prevalent in our field/applications? BMS?

3 | MODELLING FRAMEWORK

In this section, we firstly present the standard mixture cure model for a single event of interest. Then we given our general modelling framework before focusing on the application of interest. The model improves the typical approach used by linking multiple event-type modules using a hierarchical cure fraction. Throughout we refer to our motivating example of jointly modelling PFS and OS.

There are two main approaches for formulating the MCM likelihood. The first is to consider the group membership of an individual to either cured or uncured as missing data [ref]. Then the likelihood is defined in terms of the complete data and the augmented data are estimated along with the other parameters in the model, commonly using an EM algorithm. The second approach is not to be concerned with the individual classification but rather estimate a single cure fraction for the whole sample. This is the method used in this paper.

Assume that some patient-level time to event data are collected from a trial on $i = 1, \dots, n$ individuals who are given treatments $j = 1, \dots, m_i$. Denote T_{ij} to be the non-negative event times for the ith individual and treatment j. Observe a sample of event times for some parametric distribution

$$T \mid \theta \sim p(t \mid \theta).$$

Denote C_{ij} a censoring time for the ith individual and treatment j. Then define the censoring indicator $\delta_{ij} = I(T_{ij} < C_{ij})$ and variable denoting the observed survival time $t_{ij} = \min(T_{ij}, C_{ij})$. Either $\delta_{ij} = 1$ or 0 denoting an event or censored observation at t_{ij}. The observed data on the ith individual and treatment j are thus $\mathcal{O}_{ij} = (t_{ij}, \delta_{ij}, \boldsymbol{x}_{ij}),\ i = 1, \dots, N,\ j = 1, \dots, m_i \leq M$, where $\boldsymbol{x}_{ij} = (x_{ij1}, \dots, x_{ijS})$ is the individual covariate vector. Covariates may be experimental (e.g. treatment assignment) or prognostic factors.

3.1 | The standard mixture cure model

We believe the individuals to be a mixture of two distinct groups but do not observe to which group each individual belongs. Thus, the full set of model parameters consist of two separate groups of parameters $\theta = (\alpha, \mu)$. The standard mixture cure model (MCM) (sometimes called a long-term survival model) has such a set-up and is the basis of the proposed model. The MCM is a type of cure model where survival is modelled as a mixture of those who are cured and those who are not. In the simplest case, this consists of those who will never experience the event of interest and those who remain at risk of the event. The more general case adopted in this paper is when cured does not mean that an individual will never experience the event of interest but the chance of doing so reverts to a general or background population probability e.g. all-cause mortality. The combined survival for a population with a cure fraction can be written as follows

$$S(t \mid \theta, x) = S_b(t \mid \alpha, x)[\pi(x) + (1 - \pi(x))S_u(t \mid \mu, x)], \tag{1}$$

where $S(t) = 1 - \int_0^t p(s \mid \theta)\mathrm{d}s$ denotes survival at time t, $S_b(t \mid \alpha, x)$ is a function of the background mortality at time t conditional on covariates x, $\pi(x)$ denotes the probability of being cured conditional on covariates x, and $S_u(t \mid \mu, x)$ is a function of the (excess) mortality due to cancer at time t conditional on covariates x.

The survival distributions parameters α and μ comprise of a *rate* λ_α, λ_μ and a set of *ancillary* parameters ϕ_α and ϕ_μ. We can model the rate parameter using a generalised linear structure, e.g.

$$g(\lambda_i) = \beta_0^\lambda + \beta_1^\lambda x_{age,i} \; [+ \ldots],$$

where β_0^λ represents an intercept and β_1^λ the coefficients for an individual's age x_{age}. The $[+\ldots]$ term indicates additional covariates. All treatments are included in the same model with the following fixed effect model for the cure fraction

$$\text{logit}(\pi_j) = \sum_j \beta_j^\pi \text{TRT}_j \; [+ \ldots] \tag{2}$$

where β_j^π are the regression coefficients quantifying the impact of treatment TRT_j. The $[+\ldots]$ term could include a frailty term in the model. *Also, this would be naturally considered a "structured"/random effect in a Bayesian context - worth expanding on that?...*

The treatment identifier is included in the covariate vector x. We will assume without loss of generalisability that all individuals have all treatments.

In many cure fraction models it is common to have $S_b(t, x) = 1$ which is reasonable in the short-term. However, we are focused on the full life-course of an individual and so include this as a general population all-cause mortality.

3.1.1 | Background survival

We used the World Health Organisation (WHO) life tables by country for the latest year available of 2016**?** to inform the background mortality $S_b(t, x)$ in (1). The baseline hazards are the expected mortality rate for each patient at the age at which they experience the event. The mortality data are country, 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 no-one lives beyond 100 years.

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 fixed and 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 and treatment j as $\hat{p}_{ij}, \hat{S}_{ij}, \hat{h}_{ij}$ for the density, survival and hazard respectively, and include these in the set of observed data $\mathcal{O}$.

This gives

$$p(t_{ij} \mid \theta, \mathcal{O}_{ij}) = \hat{S}_{ij} \hat{h}_{ij}^{\delta_{ij}} \left[\pi(x_{ij}) + (1 - \pi(x_{ij}))S_u(t_{ij} \mid, \boldsymbol{x}_{ij}) h_u(t_{ij} \mid \boldsymbol{x}_{ij})^{\delta_{ij}}\right]$$

Sharples ref re Gamma process?

3.2 | The hierarchical mixture cure model

There may be information in one sample of event times which we can use to improve the inference from another.
Follow-up time can be relatively short for some of the event types and so there is missing data due to censoring. In this case it is even more important to make use of what information we do have available from related event times, possibly with more complete data.
In particular, OS times are necessarily longer than PFS times so will be subject to more censoring. The limited data can lead to considerable uncertainty.

The event times are often correlated because they are for the same individual. In our motivating example, $t_i^{OS} = t_i^{PFS} + t_i^{PPS}$ where t_i^{PPS} is the duration in a post-progression state and can be 0.

PFS and OS need not be for the same sample of patients.

Bivariate Normal places on the OS and PFS rates could be one option for modelling the connection**?** but we will impose dependency with hyperparameters.

Figure B2 shows a graphical representation of the general modelling framework described for OS and PFS.

[**FIGURE** 2 about here.]

We extend the observed data notation and introduce index k for multiple event types of interest i.e. for the ith individual, jth treatment and kth event type $\mathcal{O} = (t_{ijk}, \delta_{ijk}, \boldsymbol{x}_{ij}),\ k = 1, \ldots, K$. In our motivating example $K = 2$ for progression-free survival and overall survival. The set of covariates used can be different for each k and π.

The joint model is

$$p(t \mid \boldsymbol{\beta}^{\boldsymbol{\pi}}, \boldsymbol{\alpha}^{\mathbf{1}}, \ldots, \boldsymbol{\alpha}^{\boldsymbol{K}}, \boldsymbol{\mu}, \mathcal{O}) = \prod_k p(t_k \mid \boldsymbol{\beta}^{\boldsymbol{\pi}}, \boldsymbol{\alpha}^{\boldsymbol{k}}, \boldsymbol{\mu}, \mathcal{O})$$

A multilevel frailty cure fraction model for the latent variable formulation was considered by [?]. A related generalisation by [?] consider a single event type but multiple co-infections with different cure fractions, similar to how we have different cure fractions for each treatment.

3.2.1 | Representing the cure fraction when there are multiple event types

In this section we consider how to construct the extended model. We consider three alternatives.

Firstly, we can model the cure fraction corresponding to each event type completely separately, assuming that they are independent.

$$\pi_k \perp \pi_{k'},\ k, k' = 1, \ldots, K.$$

Secondly, we can assume that the cure fraction is the same for all event types and so pool the samples. In our example, PFS can be used as a proxy for OS since there are fewer missing data. There may be no reason why the cure fraction should be different between different event types, especially if individuals are observed for long enough. In this case

$$\pi_k = \pi_{k'},\ k, k' = 1, \ldots, K.$$

Lastly, and the approach adopted in this paper, is a compromise between the first two approaches called partial-pooling. We propose a hierarchical structure on the cure fraction assuming exchangeability between all event types k for each treatment j with mean value from (2).

$$\text{logit}(\pi_{jk}) \sim \text{N}(\pi_j, \sigma_j^2),\ k = 1, \ldots, K.$$

3.2.2 | Prior specification for OS, PFS model

Where possible the choice of priors should be informed by expert opinion and prior results. We specify vague priors on log-scale for the coefficients of the OS and PFS rates $\log(\lambda_{OS}), \log(\lambda_{PFS})$. Centering the ages, the baseline is a vague prior $\beta_0^{PFS} \sim \text{N}(0, 100)$, $\beta_0^{OS} \sim \text{N}(0, 100)$ and for age $\beta_{age}^{PFS} \sim \text{N}(0, 100)$, $\beta_{age}^{OS} \sim \text{N}(0, 100)$. Alternatively, we could specify the cure fraction directly using a a Beta prior distribution $\pi \sim \text{Beta}(a, b)$. The parameters can be obtained via transformation of mean and standard deviation to allow a more natural scale for elicitation.

The hierarchical cure fraction is modelled as a fixed effect linear regression with logistic link. Each treatment coefficient prior is $\beta_j^{\pi} \sim \text{N}(-0.1, 0.4)$. This corresponds to a prior mean cure fraction of just under 0.5, and 10% and 1% chance of exceeding 0.6 and 0.7 respectively.

The random effect variance on the global cure fraction is a noninformative (folded) half-Normal [?] with $\sigma^2 \sim \text{N}(0, 2.5)\boldsymbol{I}(0,)$, where $\text{N}(0, 2.5)\boldsymbol{I}(0,)$ denotes a normal distribution truncated at mean 0. Alternatively, a noninformative uniform prior density on standard deviation parameters σ is expected to generally work well when $J > 5$. For small hierarchical variance a Gamma(2, 0.1) may be preferred [?].

The background is taken directly from the WHO life-tables. However, it is believed that the patients in the trial have a worse background survival than the average individual in the general population. Using only the complete responders in the sample, who are clinically confirmed as cured, it is possible to obtain a posterior distribution for a hazard ratio between these and a WHO estimate baseline. This can serve as a prior in the main model in a two-step approach. Alternatively, prior belief could be defined explicitly using expert knowledge. This could be elicited directly for the hazard ratio or on a natural scale, such as mean life time, and transformed.

3.2.3 | Performance measures

We evaluated the goodness of fit of the models using out-of-sample predictions estimated with the widely applicable information criterion (WAIC) and leave-one-out cross-validation (LOO) [?]. These have various advantages to more common AIC and DIC and are easily obtained using the posterior sample of the Stan output.

The set of parameters in the hierarchical model (and their dimensions) are $\boldsymbol{\beta}^{\boldsymbol{\pi}}$ (3), $\boldsymbol{\sigma}$ (3), $\boldsymbol{\beta}^{\boldsymbol{\lambda}}_{\mathbf{PFS}}$ (2), $\boldsymbol{\beta}^{\boldsymbol{\lambda}}_{\mathbf{OS}}$ (2), $\boldsymbol{\phi}_{\mathbf{OS}}$ (1), $\boldsymbol{\phi}_{\mathbf{PFS}}$ (1), $\boldsymbol{\pi}_{\mathbf{OS}}$ (3), $\boldsymbol{\pi}_{\mathbf{PFS}}$ (3). Depending on the number of ancillary parameters the total is 16, 17 or 18.

In order to investigate potential decrease in the degree of uncertainty of the estimates as a result of borrowing of information across event-types, we calculated ratios of the width of the 95% CrIs. Define as the ratio of the widths of the CrIs of π_{os} and π_{pfs} from the hierarchical model to the width of the CrIs of the separate model.

4 | APPLICATION

In this section we apply the models from Section 3.2 to the Checkmate 067 dataset. The hierarchical and separate models are fit for Exponential, Weibull, Gompertz, Log-Normal, Log-Logistic to the OS and PFS event times.

4.1 | Overall results

The global cure fraction posterior is wide which is understandable since there are only two exchangeable points providing limited information (OS and PFS). However, despite this there is influence pulling the central location away from the prior downwards for *ipilimumab* and upwards for the dual treatments.

The mixed survival curves plateau to the background mortality survival from when the uncured survival curves meet the x-axis so since this is user-supplied then there is no additional insight. This is particularly early for PFS and the log-Normal distribution.

4.2 | Results of the model performance assessment

[**TABLE** 1 about here.]

[**FIGURE** 3 about here.]

[**FIGURE** 4 about here.]

[**FIGURE** 5 about here.]

[**FIGURE** 6 about here.]

4.3 | Different sample sizes due to cut-point censoring

Consider that we do not have access to the complete study follow-up data set. We shall select sensible cut-point times for demonstration purposes at 12 and 30 months. One of the benefit of adopting a hierarchical Bayesian approach is robustness to smaller sample sizes due to prior information and borrowing of strength.

Stabilisation of the cure fraction Global cure fraction is stable but uncertain

For separate PFS and OS analysis there are conflicting results.

DIC values for 12 and 30 months are...

For the separate models, for 12 months the exp/exp underestimates the OS cure fraction and then approaches the true value from below for more data. Whereas the OS cure fraction for log-Normal/log-Normal overestimates the cure fraction 12 months and then approaches the true value (for the complete data set) from above for more data. These are clearly two very different behaviours but we would not know which is correct until later on in the trial. Conversely, for the hierarchical model the cure fraction estimate is more stable about the true value.

Figures B7 and B8 show forest plots of the cure fraction estimate for 12, 30 and all months, the exp/exp and log-Normal/log-Normal models and for the separate and hierarchical models respectively.

Figures B9 and B10 show the survival curves for 12 months of follow-up data for the the exp/exp and log-Normal/log-Normal models and for the separate and hierarchical models respectively.

We can see that in Figure B7 for the exp-exp in the OS survival curves for the separate model there is more uncertainty than in the equivalent plot for the hierarchical model in B8.

Also, Figure B7 for the log-Normal/log-Normal in the OS survival curves for separate and hierarchical models it over-estimates the cure fraction. It appears to revert to the background survival at the time of the last observed data point.

The forest plots demonstrate the same behaviour. Figure B9 Figure B10

[**FIGURE** 7 about here.]

[**FIGURE** 8 about here.]

[**FIGURE** 9 about here.]

[**FIGURE** 10 about here.]

The ratio of the CrIs of the separate and hierarchical cure fractions didn't show any significant difference for the complete data set. For the cut-point samples...

5 | DISCUSSION

Our approach allows for the incorporation of sensible information. Stabilise the inference on the cure fraction but restricting/-constraining it values via the prior distributions.

In comparison to some previous frequentist approaches, our framework is able to include contextual information about the parameters of the model via the prior distributions. For this analysis we chose uninformative priors. This allows the data to dominate the posterior even if the dataset is relatively small. In practice there may be additional information from clinical experience or previous data that can better inform the prior selection. The priors can be specified on the natural scale to aid elicitation and interpretation. For example, we may have clinical knowledge that a drug treatment could generate a cure fraction greater than, say, 30%. If a (cancer) patient enters the study at 60 years old then there is a high degree of certainty that they will not live for, say, over 40 years. In the case of sparse data e.g. when there is a large amount of missing data perhaps due to right censoring, then 'regularising' the inference based on available prior information could make a significant and important difference.

How to elicit this prior information in practice may not be simple and a formal protocol should be adopted[?].

The regulators may value the fact that a hierarchical model can embed some form of prior knowledge to account for scepticism. This will prevent the cure rates to be taken at face value, particularly with limited data. The hierarchical model seems to give a more precise estimate of the cure rate, even with earlier data cuts (especially for the Exponential model). The hierarchical model is more aligned, which means possibly you have more reliable estimates early on, which is a good argument to complement modelling based on earlier cuts. Earlier cuts can lead to quicker introduction in the market and reimbursement decisions.

Another way to think about the cut-point data is to think about what time should we choose to halt the trial and end the data collection which would maximise some expected overall utility. That is what is the optimal stopping time τ^* for the total utility V.

$$\mathbb{E}V_{\tau^*} = \sup_{\tau} \mathbb{E}V_{\tau}$$

The stopping utility may include benefits such as smaller costs, freed-up time from the trial for patients and those running the study, quicker time to market for a drug, avoidance of side-effects etc. The utility would also depend on estimating 'correct' parameters values from the available data i.e. that they are within some acceptable threshold of the true values. This could be directly on the parameters of interest or an order statistics for which drug is deemed the preferable.

This may be framed in terms of value of information (VoI) [ref]. For example, in the case given above, what is the additional value of obtaining further data after 12 and 30 months.

In practice, deriving a set of heuristics may be better than formalising a general rule. A discrete set of cut-point may be used, similar to our selection of 12 and 30 months, rather than considering the whole real line.

An important use of the output from this work is in cost-effectiveness analysis. The survival times are use directly as an effectiveness measure or in a function to obtain e.g. QALYS or QALE. For unstable cure fraction estimates, such as was the case for a small sample of data for OS in the separate model, would give very different QALY estimates and thus cost-effectiveness statistics and potentially differing optimal decisions. The hierarchical model, which exhibited more stable estimation, would be more reliable in this situation.

how should the complete responders be included in the main model? If they enter via the background adjustment why double-count them? Landmarking?

It may be that there is excessive borrowing of information between cure fractions when the event types are distinctly different. In our example, this is not necessary since there are only two event types but for a larger number this may be appropriate. Partial-exchangeability can be used instead in this case[?].

In terms of an outcome of interest and a surrogate end-point Papanikos (2020)[?] modelled their relationship using a Bivariate Normal distribution. We could have done something similar for the OS and PFS event times, especially since the latter is a lower bound for the former.

model averaging using all distributions and weights from LOO/WAIC?

extensions of the implementation could be to generalise to have different distributions for each treatment in a single model fit, rather than at the moment where they are assumed to be the same for all treatments.

ACKNOWLEDGEMENTS

This is acknowledgement text.

Author contributions

This is an author contribution text.

Financial disclosure

None reported.

Conflict of interest

The authors declare no potential conflict of interests.

How to cite this article: Green N., and G. Baio (2021), , *2017;00:1–6*.

APPENDIX

A POSTERIOR PARAMETER STATISTICS

[**TABLE** 2 about here.]

[**TABLE** 3 about here.]

[**TABLE** 4 about here.]

[**TABLE** 5 about here.]

[**TABLE** 6 about here.]

[**TABLE** 7 about here.]

[**TABLE** 8 about here.]

[**TABLE** 9 about here.]

[**TABLE** 10 about here.]

[**TABLE** 11 about here.]

[**TABLE** 12 about here.]

[**TABLE** 13 about here.]

[**TABLE** 14 about here.]

[**TABLE** 15 about here.]

B POSTERIOR PREDICTIVE CHECKS

[**FIGURE** 11 about here.]

References

1. Amico M, Van Keilegom I. Cure Models in Survival Analysis. *Annual Review of Statistics and Its Application* 2018; 5: 311–342. doi: 10.1146/annurev-statistics-031017-100101

2. Yu X, De Angelis R, Andersson T, Lambert P, O'Connell D, Dickman P. Estimating the proportion cured of cancer: Some practical advice for users. *Cancer Epidemiology* 2013; 37(6): 836–842. doi: 10.1016/j.canep.2013.08.014

3. Demiris N, Sharples LD. Bayesian evidence synthesis to extrapolate survival estimates in cost-effectiveness studies. *Statistics in Medicine* 2006; 25(11): 1960–1975. doi: 10.1002/sim.2366

4. Jackson CH, Sharples LD, Thompson SG. Survival models in health economic evaluations: balancing fit and parsimony to improve prediction.. *The international journal of biostatistics* 2010; 6(1): Article 34. doi: 10.2202/1557-4679.1269

5. Lai X, Yau KKW. Multilevel Mixture Cure Models with Random Effects. *Biometrical Journal* 2009; 51(3): 456–466. doi: 10.1002/bimj.200800222

6. Lai X, Yau KKW. Long-term survivor model with bivariate random effects: Applications to bone marrow transplant and carcinoma study data. *Statistics in Medicine* 2008; 27(27): 5692–5708. doi: 10.1002/sim.3404

7. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. *Sheffield: Report by the Decision Support Unit* 2011(0).

8. Carpenter B, Gelman A, Hoffman MD, et al. Stan: A probabilistic programming language. *Journal of statistical software* 2017; 76(1).

9. R Core Team . *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; Vienna, Austria: 2020.

10. WHO . Life tables by country. 2020. http://apps.who.int/gho/data/ (accessed on 02/12/20).

11. Tan SH, Abrams KR, Bujkiewicz S. Bayesian Multiparameter Evidence Synthesis to Inform Decision Making: A Case Study in Metastatic Hormone-Refractory Prostate Cancer. *Medical Decision Making* 2018; 38(7): 834–848. doi: 10.1177/0272989X18788537

12. Tawiah R, McLachlan GJ, Ng SK. Mixture cure models with time-varying and multilevel frailties for recurrent event data. *Statistical Methods in Medical Research* 2020; 29(5): 1368–1385. doi: 10.1177/0962280219859377

13. Balogun OS, Gao XZ, Jolayemi ET, Olaleye SA. Generalized cure rate model for infectious diseases with possible co-infections. *PLoS ONE* 2020; 15(9 September): 1–16. doi: 10.1371/journal.pone.0239003

14. Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 2006; 1(3): 515-–533.

15. Chung Y, Rabe-Hesketh S, Dorie V, Gelman A, Liu J. A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika* 2013(4): 685-–709.

16. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 2017; 27(5): 1413–1432. doi: 10.1007/s11222-016-9696-4

17. O'Hagan A. Expert Knowledge Elicitation: Subjective but Scientific. *American Statistician* 2019; 73(sup1): 69–81. doi: 10.1080/00031305.2018.1518265

18. Neuenschwander B, Wandel S, Roychoudhury S, Bailey S. Robust exchangeability designs for early phase clinical trials with multiple strata. *Pharmaceutical Statistics* 2016; 15(2): 123–134. doi: 10.1002/pst.1730

19. Papanikos T, Thompson JR, Abrams KR, et al. Bayesian hierarchical meta-analytic methods for modeling surrogate relationships that vary across treatment classes using aggregate data. *Statistics in Medicine* 2020; 39(8): 1103–1124. doi: 10.1002/sim.8465

List of Figures

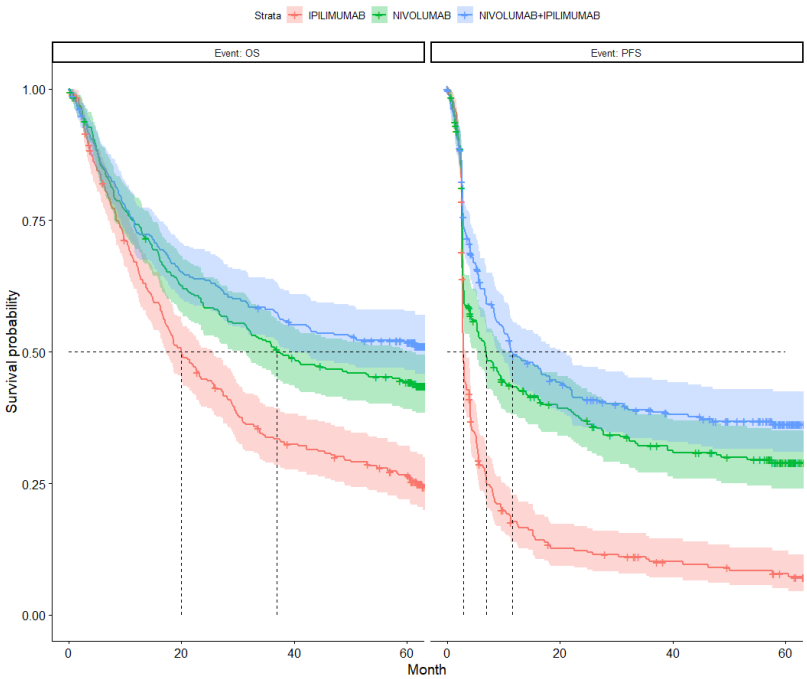


FIGURE B1 Kaplan-Meier curves of OS and PFS times with 95% CI for the Checkmate 067 study data and *Ipilimumab*, *Nivolumab* and combination treatments. Median survival times are shown with dashed lines.

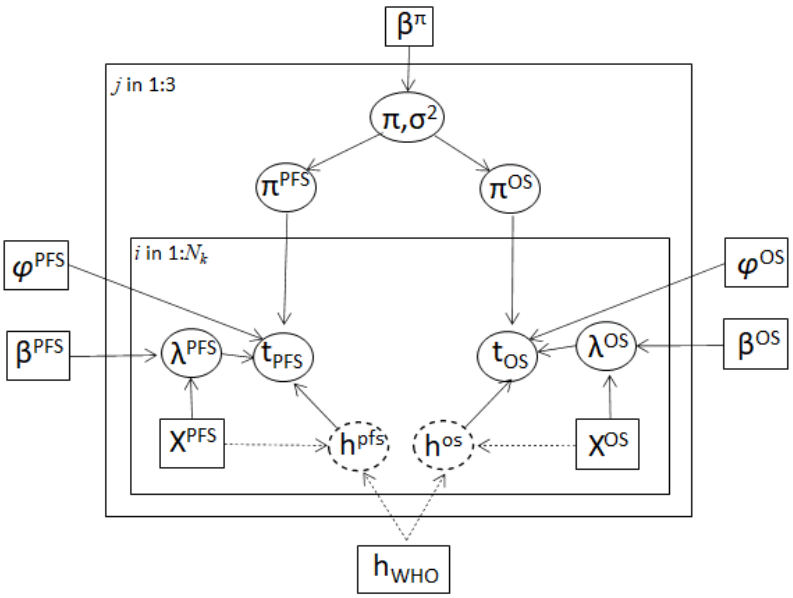


FIGURE B2 Hierarchical cure fraction DAG for PFS and OS. Solid lines represent stochastic and dashed lines deterministic relationships, respectively. Cured patients have fixed hazards taken from WHO life-tables. The distribution of times for uncured patient regresses on covariates for the rate parameter and ϕ is the set of ancillary parameters. Censoring has been omitted for brevity.

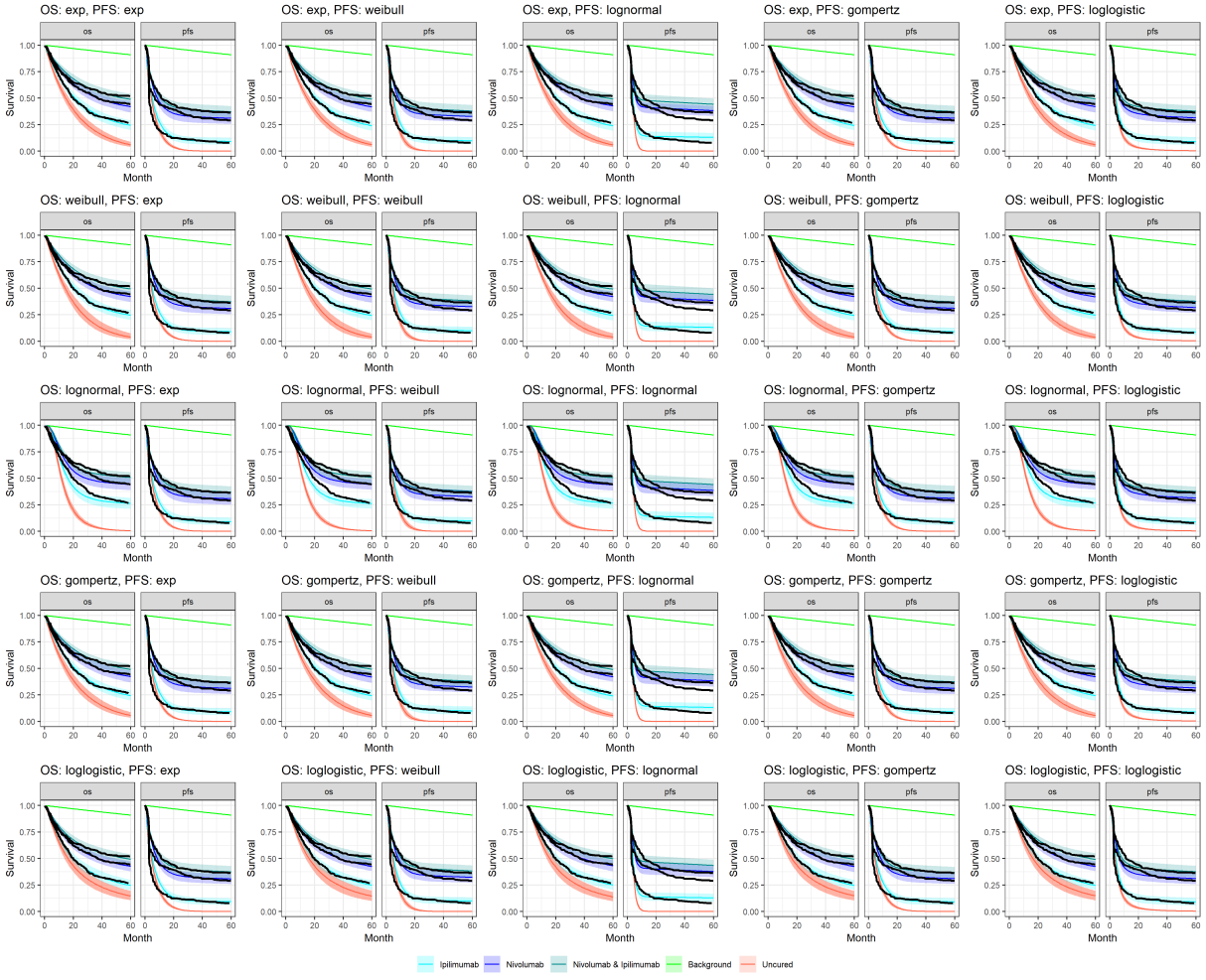


FIGURE B3 Hierarchical model posterior survival curves for exponential, weibull, gompertz, log-logistic and log-Normal uncured fraction for OS and PFS events and *ipilimumab*, *nivolumab* and combination treatments. The black lines show the Kaplan-Meier curves.

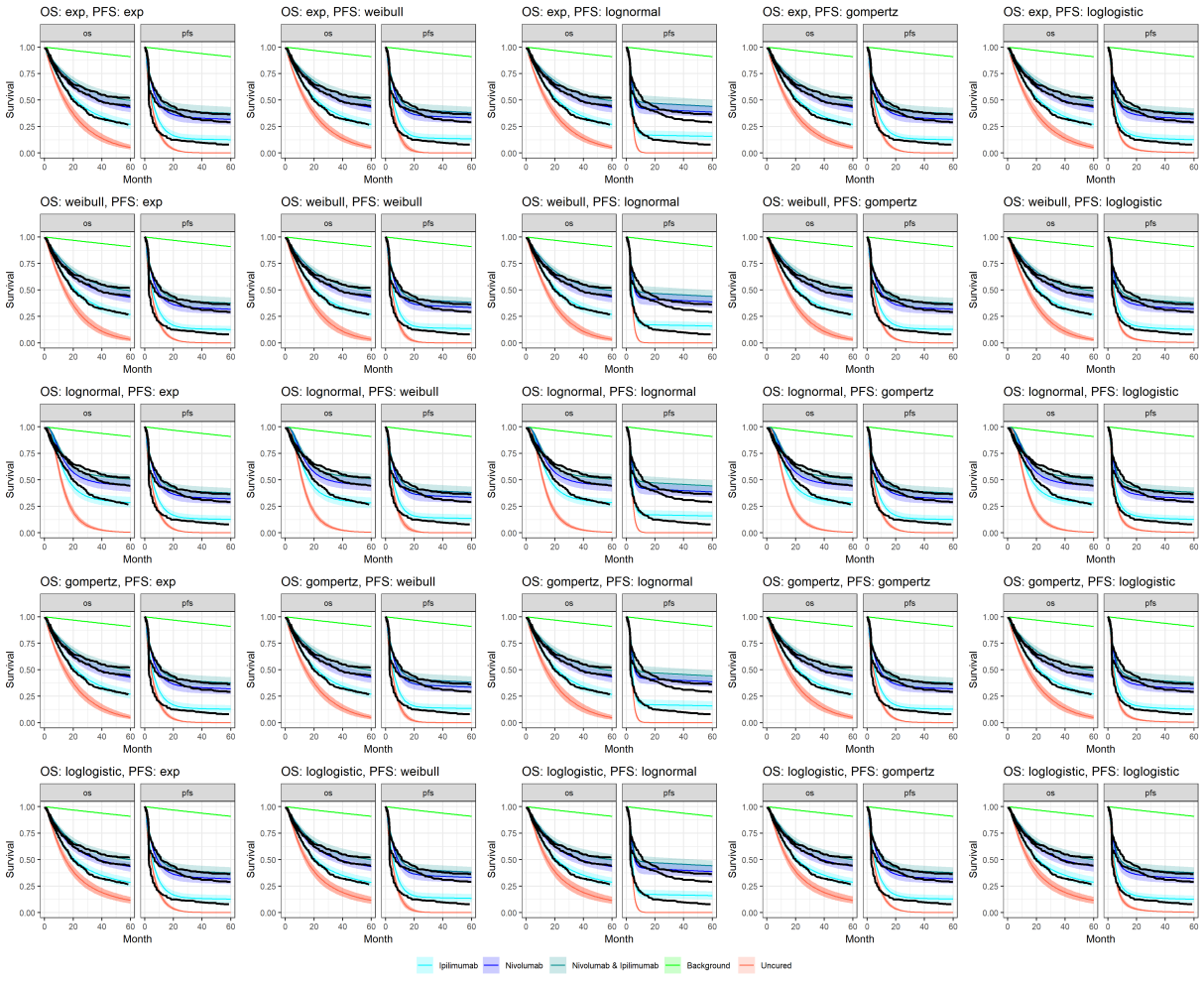


FIGURE B4 Separate model posterior survival curves for exponential, weibull, gompertz, log-logistic and log-Normal uncured fraction for OS and PFS events *ipilimumab*, *nivolumab* and combination treatments. The black lines show the Kaplan-Meier curves.

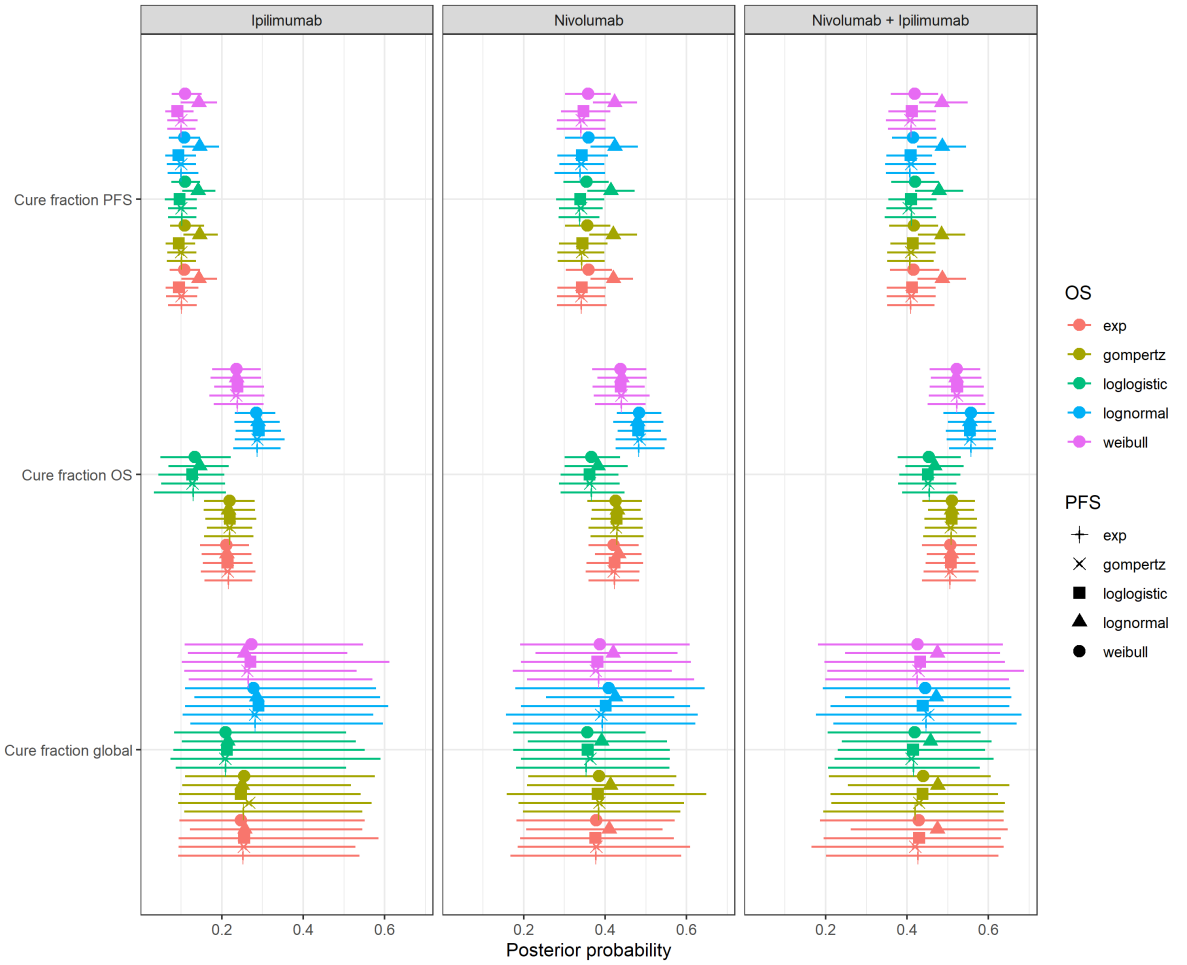


FIGURE B5 Hierarchical model posterior cure fraction forest plots with 95% credible intervals for (i) *ipilimumab* only (ii) *nivolumab* only and (iii) combination treatment.

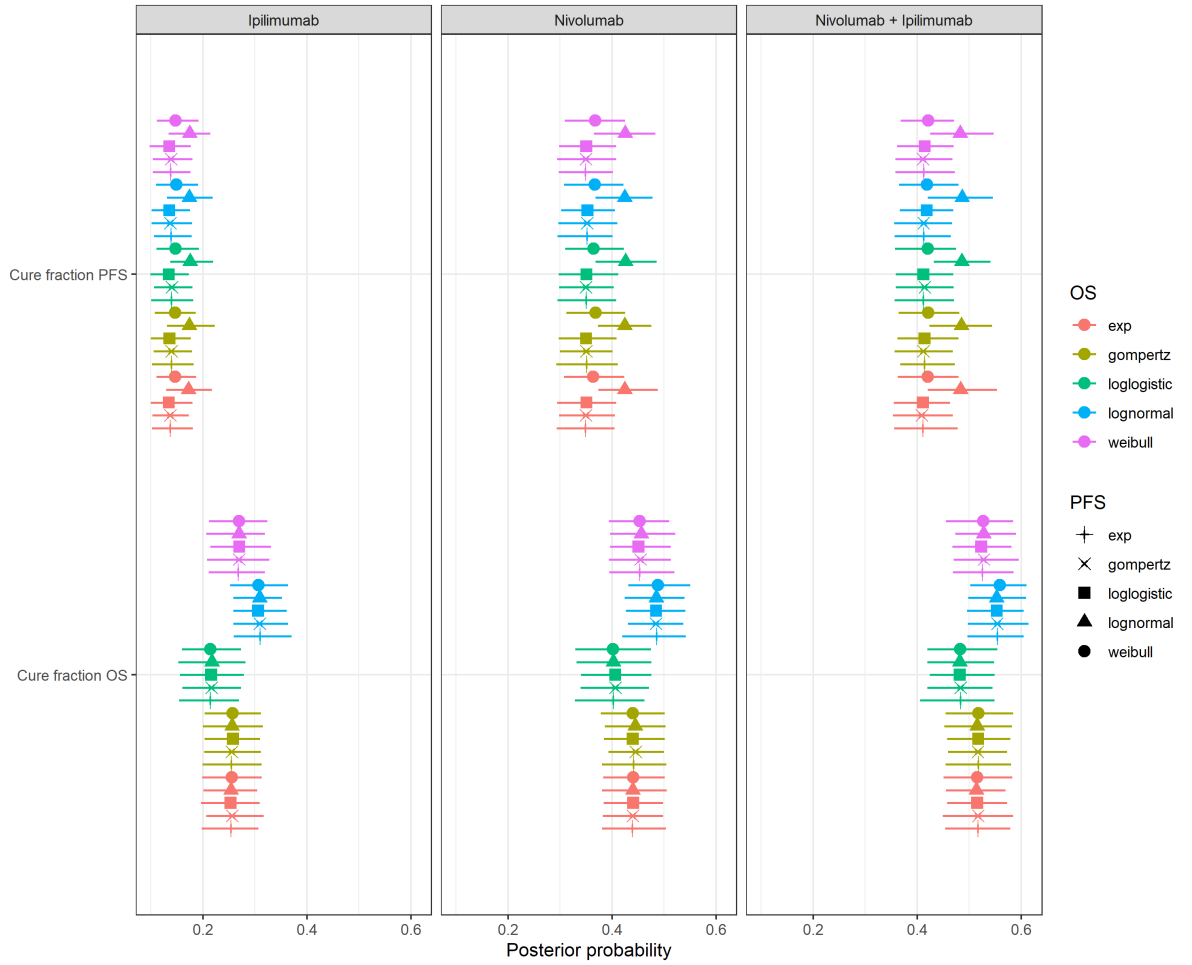


FIGURE B6 Separate model posterior cure fraction forest plots with 95% credible intervals for (i) *ipilimumab* only (ii) *nivolumab* only and (iii) combination treatment.

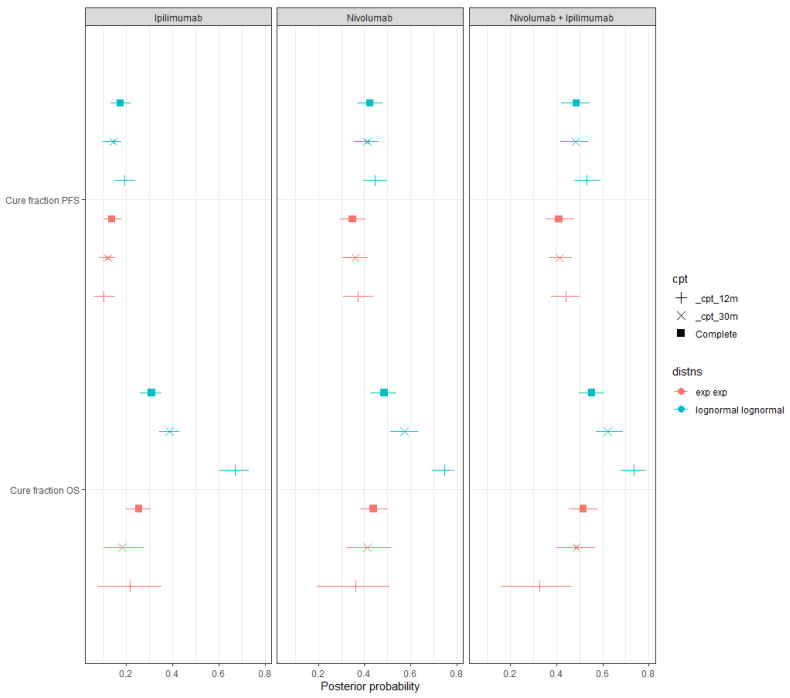


FIGURE B7 Posterior cure fraction forest plot for study data censored at 12 months, 30 months and the complete data set using the separate model.

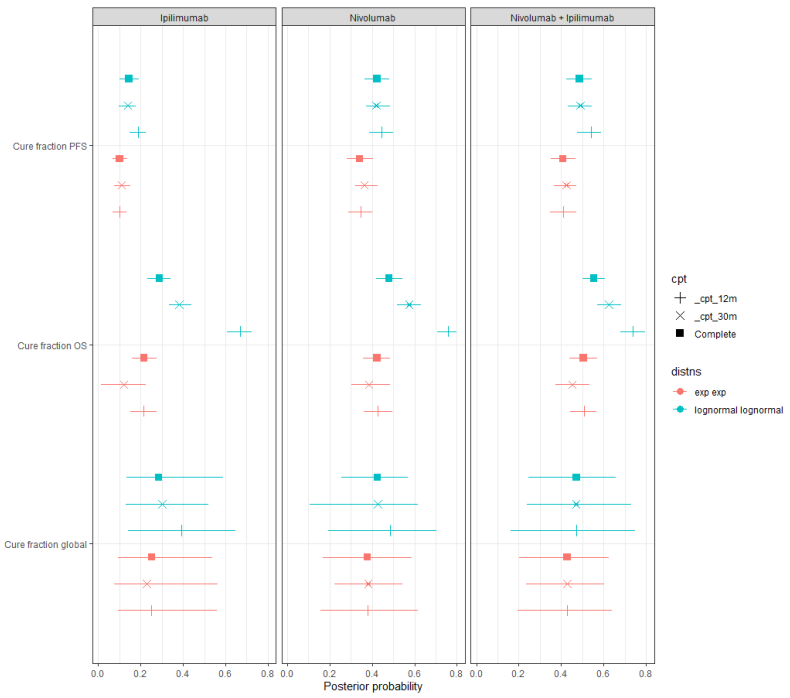


FIGURE B8 Posterior cure fraction forest plot for study data censored at 12 months, 30 months and the complete data set using the hierarchical model.

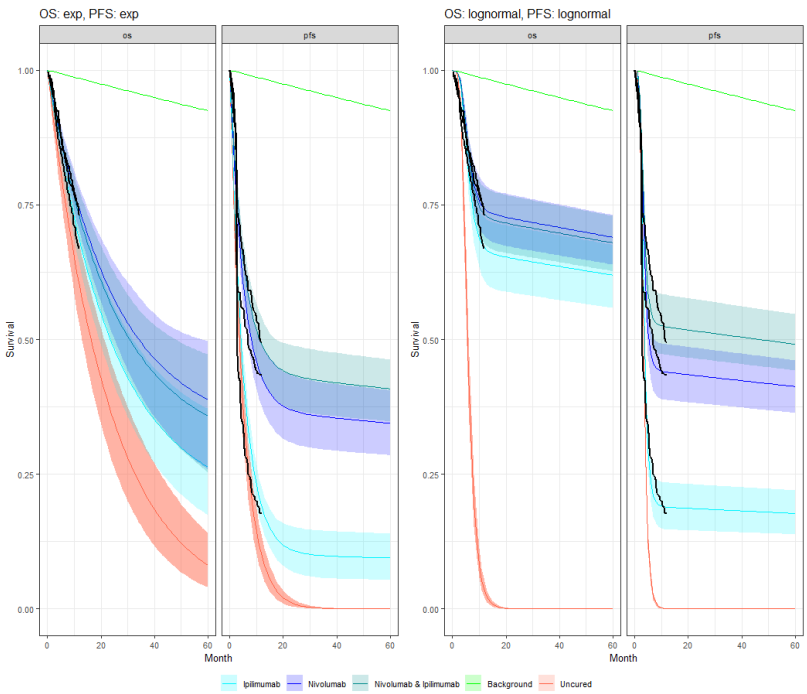


FIGURE B9 Separate model posterior survival curves for exponential/exponential and log-Normal/log-Normal uncured fraction for OS/PFS events and censored times at 12 months. The black lines show the observed data Kaplan-Meier curves.

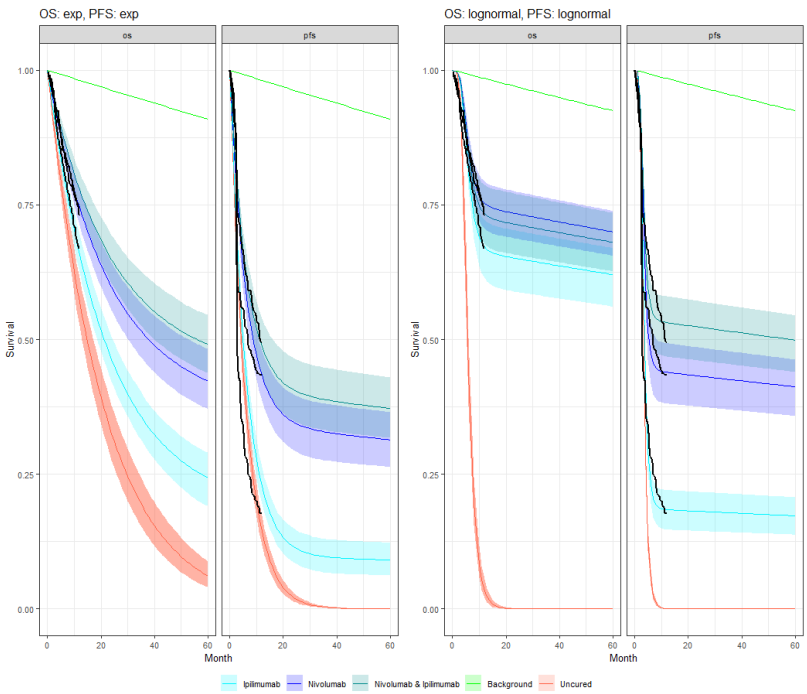
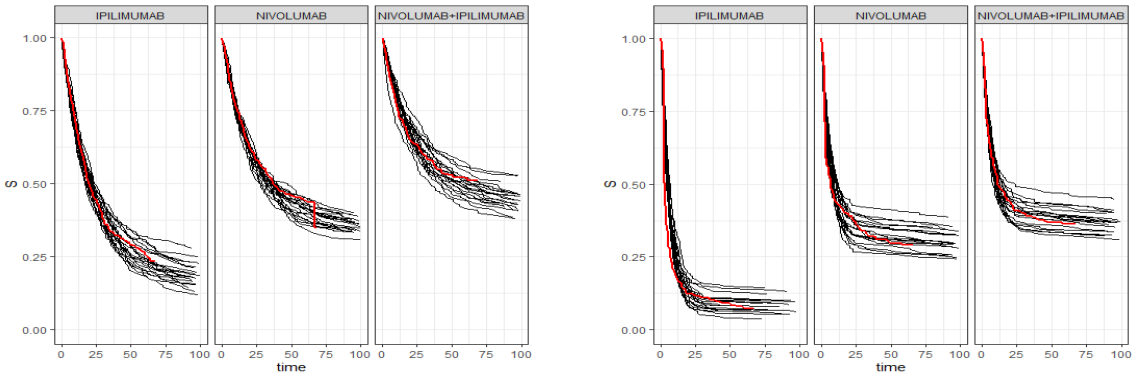
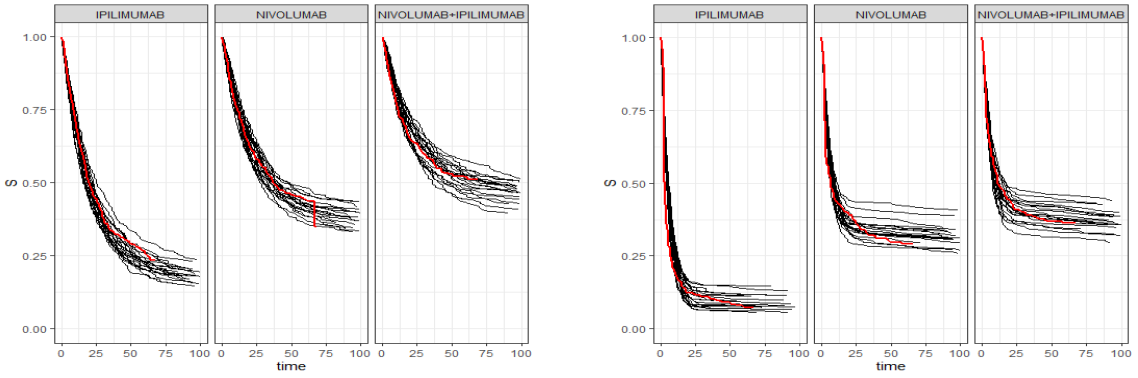


FIGURE B10 Hierarchical model posterior survival curves for exponential/exponential and log-Normal/log-Normal uncured fraction for OS/PFS events and censored times at 12 months. The black lines show the observed data Kaplan-Meier curves.



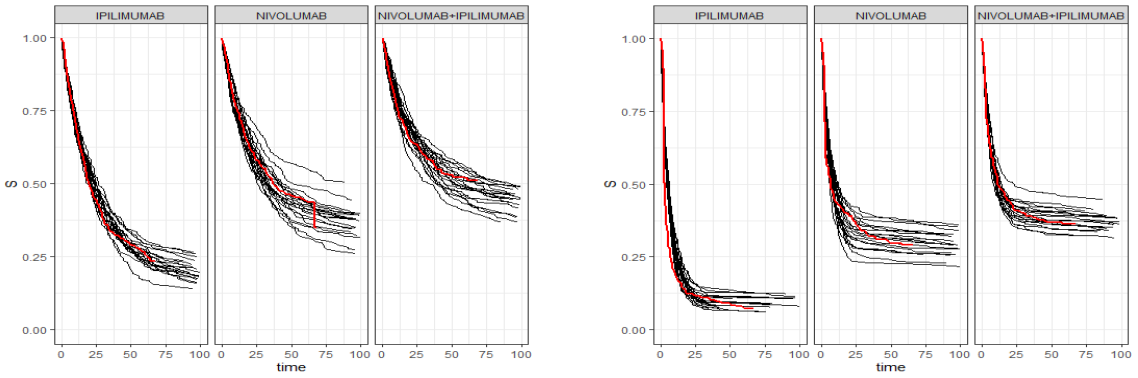
(a) OS data **(b)** PFS data

FIGURE B11 Posterior predictive values Kaplan-Meier curves for exponential OS and exponential PFS distributions with the hierarchical model. Black lines represent outcomes for 20 predicted cohorts and the red line is for the observed data.



(a) OS data **(b)** PFS data

FIGURE B12 Posterior predictive values Kaplan-Meier curves for weibull OS and weibull PFS distributions with the hierarchical model. Black lines represent outcomes for 20 predicted cohorts and the red line is for the observed data.



(a) OS data **(b)** PFS data

FIGURE B13 Posterior predictive values Kaplan-Meier curves for gompertz OS and gompertz PFS distributions with the hierarchical model. Black lines represent outcomes for 20 predicted cohorts and the red line is for the observed data.

List of Tables

TABLE B1 WAIC statistics for hierarchical and separate models, and all distributions.

		Hierarchical						Separate				
OS distn	**PFS distn**	**ELPD**[†]	**SE**	$p_D^{\ddagger}$	**SE**	**WAIC**	**SE**	**ELPD**	**SE**	p_D	**SE**	**WAIC**
exp	exp	-5064.70	80.37	11.13	0.46	10129.40	160.74	-5066.38	81.40	10.46	0.43	10132.76
exp	gompertz	-5065.64	80.44	11.80	0.49	10131.29	160.88	-5066.68	81.46	10.38	0.46	10133.37
exp	loglogistic	-4970.34	80.72	11.78	0.33	9940.68	161.45	-4971.73	81.67	10.36	0.28	9943.46
exp	lognormal	-4730.47	92.79	18.04	0.91	9460.94	185.59	-4732.41	93.18	18.56	0.99	9464.83
exp	weibull	-5062.99	81.69	14.63	0.80	10125.97	163.38	-5064.87	82.86	14.39	0.86	10129.74
gompertz	exp	-5064.99	80.41	11.45	0.45	10129.99	160.82	-5066.72	81.44	10.51	0.45	10133.43
gompertz	gompertz	-5065.61	80.44	11.66	0.48	10131.22	160.88	-5066.76	81.49	10.19	0.44	10133.53
gompertz	loglogistic	-4970.42	80.79	11.68	0.34	9940.84	161.59	-4972.24	81.70	10.70	0.28	9944.47
gompertz	lognormal	-4730.92	92.64	18.31	0.95	9461.84	185.29	-4732.23	93.30	18.01	0.95	9464.46
gompertz	weibull	-5063.81	81.74	15.45	0.85	10127.61	163.48	-5064.46	82.92	13.50	0.78	10128.91
loglogistic	exp	-5059.86	80.47	11.29	0.42	10119.73	160.93	-5062.51	81.89	10.79	0.40	10125.02
loglogistic	gompertz	-5060.37	80.40	11.64	0.45	10120.74	160.80	-5062.64	82.01	10.26	0.41	10125.27
loglogistic	loglogistic	-4966.22	80.75	12.77	0.35	9932.44	161.50	-4968.45	82.36	11.63	0.28	9936.89
loglogistic	lognormal	-4726.57	93.17	19.43	0.91	9453.13	186.34	-4728.25	94.20	18.18	0.92	9456.51
loglogistic	weibull	-5058.17	81.84	14.80	0.74	10116.33	163.67	-5060.18	83.31	13.79	0.69	10120.37
lognormal	exp	-4972.55	84.58	16.35	0.61	9945.09	169.15	-4973.43	85.26	14.63	0.55	9946.86
lognormal	gompertz	-4973.16	84.69	16.50	0.56	9946.32	169.38	-4973.48	85.22	14.31	0.51	9946.96
lognormal	loglogistic	-4877.87	85.79	16.53	0.50	9755.74	171.59	-4879.07	86.55	15.07	0.44	9758.14
lognormal	lognormal	-4638.85	100.27	23.58	1.00	9277.71	200.55	-4638.43	100.40	21.30	0.96	9276.85
lognormal	weibull	-4970.22	86.17	19.01	0.89	9940.44	172.34	-4972.48	87.07	19.54	0.91	9944.95
weibull	exp	-5063.26	80.70	11.93	0.46	10126.53	161.40	-5064.74	81.52	11.11	0.45	10129.48
weibull	gompertz	-5064.37	80.64	12.67	0.52	10128.73	161.28	-5065.16	81.59	11.16	0.44	10130.32
weibull	lognormal	-4729.78	93.18	19.54	0.97	9459.56	186.36	-4730.93	93.78	19.04	0.96	9461.85
weibull	weibull	-5061.91	82.05	15.77	0.85	10123.81	164.11	-5063.20	83.09	14.87	0.83	10126.41

[†]ELPD: Expected log pointwise predictive density; [‡]p_D: Effective number of parameters

TABLE B2 Posterior summary statistics for hierarchical model parameters and all OS distributions with exponential PFS distribution.

	Exp/Exp[1]		Weibull/Exp		Gompertz/Exp		Log-logistic/Exp	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
$\beta_{os,0}$	-3.093	(-3.249, -2.956)	3.032	(2.903, 3.221)	-3.094	(-3.264, -2.948)	2.831	(2.659, 3.023)
$\beta_{os,1}$	-0.014	(-0.023, -0.005)	0.011	(0.003, 0.021)	-0.014	(-0.024, -0.003)	0.011	(0, 0.021)
$\beta_{pfs,0}$	-1.825	(-1.939, -1.728)	-1.828	(-1.925, -1.738)	-1.827	(-1.92, -1.73)	-1.831	(-1.921, -1.734)
$\beta_{pfs,1}$	-0.003	(-0.009, 0.004)	-0.003	(-0.009, 0.003)	-0.003	(-0.01, 0.003)	-0.003	(-0.009, 0.003)
β_1^{π}	-1.179	(-2.283, 0.153)	-1.102	(-2.009, 0.28)	-1.168	(-2.115, 0.18)	-1.447	(-2.351, 0.022)
β_2^{π}	-0.527	(-1.604, 0.351)	-0.494	(-1.337, 0.485)	-0.496	(-1.4, 0.341)	-0.625	(-1.509, 0.231)
β_3^{π}	-0.311	(-1.382, 0.508)	-0.324	(-1.392, 0.62)	-0.348	(-1.422, 0.565)	-0.354	(-1.35, 0.318)
σ_1	1.403	(0.339, 3.87)	1.486	(0.325, 4.266)	1.542	(0.302, 4.395)	1.268	(0.058, 4.274)
σ_2	0.791	(0.056, 2.834)	0.868	(0.111, 3.433)	0.846	(0.064, 2.911)	0.701	(0.012, 2.897)
σ_3	0.852	(0.091, 2.636)	0.985	(0.146, 3.444)	1.008	(0.116, 3.424)	0.703	(0.016, 2.906)
$\pi_{global,1}$	0.252	(0.093, 0.538)	0.265	(0.118, 0.57)	0.253	(0.108, 0.545)	0.209	(0.087, 0.505)
$\pi_{global,2}$	0.377	(0.167, 0.587)	0.384	(0.208, 0.619)	0.384	(0.198, 0.584)	0.354	(0.181, 0.557)
$\pi_{global,3}$	0.427	(0.201, 0.624)	0.425	(0.199, 0.65)	0.420	(0.194, 0.638)	0.416	(0.206, 0.579)
$\pi_{os,1}$	0.216	(0.158, 0.275)	0.238	(0.181, 0.302)	0.219	(0.156, 0.278)	0.130	(0.033, 0.211)
$\pi_{os,2}$	0.423	(0.359, 0.484)	0.440	(0.375, 0.499)	0.429	(0.364, 0.494)	0.366	(0.291, 0.447)
$\pi_{os,3}$	0.505	(0.437, 0.569)	0.523	(0.45, 0.593)	0.508	(0.439, 0.569)	0.455	(0.388, 0.524)
$\pi_{pfs,1}$	0.101	(0.068, 0.139)	0.099	(0.065, 0.135)	0.100	(0.065, 0.137)	0.102	(0.068, 0.137)
$\pi_{pfs,2}$	0.341	(0.282, 0.404)	0.340	(0.281, 0.401)	0.342	(0.284, 0.398)	0.338	(0.285, 0.386)
$\pi_{pfs,3}$	0.409	(0.351, 0.468)	0.410	(0.353, 0.472)	0.407	(0.351, 0.466)	0.411	(0.345, 0.472)

1OS distribution/PFS distribution.

TABLE B3 Posterior summary statistics for hierarchical model parameters and all OS distributions with weibull PFS distribution.

	Exp/Weibull[1]		Weibull/Weibull		Gompertz/Weibull		Log-logistic/Weibull	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
$\beta_{os,0}$	-3.097	(-3.247, -2.968)	3.032	(2.894, 3.198)	-3.092	(-3.243, -2.949)	2.826	(2.661, 3.011)
$\beta_{os,1}$	-0.014	(-0.023, -0.005)	0.011	(0.002, 0.02)	-0.013	(-0.023, -0.004)	0.011	(0.002, 0.022)
$\beta_{pfs,0}$	1.799	(1.705, 1.899)	1.797	(1.7, 1.899)	1.798	(1.711, 1.904)	1.802	(1.711, 1.901)
$\beta_{pfs,1}$	0.002	(-0.004, 0.007)	0.001	(-0.005, 0.007)	0.001	(-0.005, 0.008)	0.002	(-0.004, 0.007)
β_1^{π}	-1.208	(-2.243, 0.204)	-1.056	(-2.104, 0.188)	-1.159	(-2.094, 0.307)	-1.450	(-2.403, 0.021)
β_2^{π}	-0.518	(-1.498, 0.286)	-0.482	(-1.445, 0.439)	-0.487	(-1.319, 0.3)	-0.612	(-1.554, -0.001)
β_3^{π}	-0.306	(-1.473, 0.563)	-0.324	(-1.51, 0.557)	-0.259	(-1.334, 0.428)	-0.340	(-1.353, 0.327)
σ_1	1.359	(0.223, 4.202)	1.485	(0.267, 4.21)	1.335	(0.247, 3.67)	1.089	(0.036, 3.606)
σ_2	0.764	(0.026, 3.09)	0.851	(0.071, 3.113)	0.704	(0.033, 2.6)	0.639	(0.015, 2.502)
σ_3	0.888	(0.09, 2.978)	1.081	(0.125, 3.455)	0.825	(0.09, 2.732)	0.572	(0.01, 2.264)
$\pi_{global,1}$	0.247	(0.096, 0.551)	0.273	(0.109, 0.547)	0.255	(0.11, 0.576)	0.209	(0.083, 0.505)
$\pi_{global,2}$	0.378	(0.183, 0.571)	0.387	(0.191, 0.608)	0.385	(0.211, 0.574)	0.356	(0.175, 0.5)
$\pi_{global,3}$	0.429	(0.186, 0.637)	0.426	(0.181, 0.636)	0.440	(0.209, 0.605)	0.419	(0.205, 0.581)
$\pi_{os,1}$	0.211	(0.147, 0.266)	0.236	(0.177, 0.296)	0.219	(0.156, 0.281)	0.134	(0.049, 0.222)
$\pi_{os,2}$	0.421	(0.359, 0.482)	0.438	(0.369, 0.501)	0.426	(0.356, 0.49)	0.366	(0.3, 0.437)
$\pi_{os,3}$	0.506	(0.436, 0.572)	0.522	(0.456, 0.58)	0.511	(0.438, 0.567)	0.454	(0.377, 0.532)
$\pi_{pfs,1}$	0.108	(0.072, 0.146)	0.109	(0.077, 0.15)	0.109	(0.073, 0.156)	0.109	(0.076, 0.147)
$\pi_{pfs,2}$	0.359	(0.304, 0.417)	0.358	(0.301, 0.414)	0.356	(0.301, 0.413)	0.354	(0.297, 0.409)
$\pi_{pfs,3}$	0.416	(0.358, 0.479)	0.419	(0.36, 0.476)	0.417	(0.356, 0.477)	0.420	(0.361, 0.478)

1OS distribution/PFS distribution.

TABLE B4 Posterior summary statistics for hierarchical model parameters and all OS distributions with gompertz PFS distribution.

	Exp/Gompertz[1]		Weibull/Gompertz		Gompertz/Gompertz		Log-logistic/Gompertz	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
$\beta_{os,0}$	-3.094	(-3.247, -2.956)	3.036	(2.922, 3.196)	-3.091	(-3.27, -2.944)	2.844	(2.671, 3.03)
$\beta_{os,1}$	-0.014	(-0.023, -0.005)	0.011	(0.002, 0.02)	-0.013	(-0.023, -0.004)	0.011	(0.002, 0.021)
$\beta_{pfs,0}$	-1.827	(-1.929, -1.727)	-1.834	(-1.933, -1.729)	-1.831	(-1.947, -1.721)	-1.828	(-1.92, -1.726)
$\beta_{pfs,1}$	-0.003	(-0.01, 0.004)	-0.003	(-0.009, 0.004)	-0.003	(-0.009, 0.003)	-0.003	(-0.01, 0.003)
β_1^{π}	-1.171	(-2.273, 0.111)	-1.120	(-2.117, 0.125)	-1.105	(-2.281, 0.272)	-1.465	(-2.526, 0.364)
β_2^{π}	-0.521	(-1.482, 0.441)	-0.526	(-1.562, 0.259)	-0.490	(-1.468, 0.38)	-0.579	(-1.427, 0.235)
β_3^{π}	-0.350	(-1.62, 0.563)	-0.311	(-1.36, 0.785)	-0.299	(-1.299, 0.579)	-0.375	(-1.256, 0.46)
σ_1	1.362	(0.3, 4.001)	1.481	(0.38, 4.043)	1.548	(0.324, 3.606)	1.142	(0.034, 3.456)
σ_2	0.810	(0.06, 3.291)	0.960	(0.092, 3.207)	0.933	(0.088, 3.125)	0.607	(0.013, 2.766)
σ_3	0.973	(0.081, 3.63)	0.990	(0.151, 3.144)	0.912	(0.07, 3.307)	0.692	(0.033, 3.032)
$\pi_{global,1}$	0.254	(0.093, 0.528)	0.262	(0.107, 0.531)	0.267	(0.093, 0.568)	0.208	(0.074, 0.59)
$\pi_{global,2}$	0.378	(0.185, 0.608)	0.377	(0.173, 0.564)	0.386	(0.187, 0.594)	0.363	(0.194, 0.559)
$\pi_{global,3}$	0.420	(0.165, 0.637)	0.428	(0.204, 0.687)	0.429	(0.214, 0.641)	0.411	(0.222, 0.613)
$\pi_{os,1}$	0.214	(0.148, 0.282)	0.235	(0.169, 0.304)	0.219	(0.163, 0.275)	0.128	(0.051, 0.208)
$\pi_{os,2}$	0.421	(0.352, 0.485)	0.439	(0.372, 0.509)	0.426	(0.36, 0.492)	0.363	(0.287, 0.436)
$\pi_{os,3}$	0.507	(0.441, 0.576)	0.522	(0.453, 0.588)	0.507	(0.444, 0.571)	0.452	(0.378, 0.521)
$\pi_{pfs,1}$	0.101	(0.063, 0.139)	0.099	(0.066, 0.14)	0.100	(0.066, 0.137)	0.100	(0.068, 0.139)
$\pi_{pfs,2}$	0.341	(0.282, 0.4)	0.342	(0.282, 0.401)	0.343	(0.283, 0.397)	0.341	(0.286, 0.394)
$\pi_{pfs,3}$	0.411	(0.35, 0.471)	0.408	(0.349, 0.469)	0.410	(0.351, 0.47)	0.404	(0.349, 0.462)

1OS distribution/PFS distribution.

TABLE B5 Posterior summary statistics for hierarchical model parameters and all OS distributions with log-logistic PFS distribution.

	Exp/Log-logistic[1]		Weibull/Log-logistic		Gompertz/Log-logistic		Log-logistic/Log-logistic	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
$\beta_{os,0}$	-3.093	(-3.258, -2.954)	3.025	(2.915, 3.185)	-3.090	(-3.241, -2.923)	2.842	(2.659, 3.034)
$\beta_{os,1}$	-0.014	(-0.022, -0.005)	0.011	(0.003, 0.02)	-0.013	(-0.023, -0.004)	0.011	(0.002, 0.021)
$\beta_{pfs,0}$	1.334	(1.257, 1.41)	1.337	(1.269, 1.406)	1.334	(1.264, 1.406)	1.340	(1.259, 1.413)
$\beta_{pfs,1}$	0.002	(-0.002, 0.007)	0.002	(-0.002, 0.008)	0.002	(-0.002, 0.007)	0.002	(-0.002, 0.007)
β_1^{π}	-1.170	(-2.267, 0.343)	-1.085	(-2.176, 0.456)	-1.209	(-2.253, 0.164)	-1.440	(-2.431, 0.204)
β_2^{π}	-0.529	(-1.444, 0.278)	-0.508	(-1.428, 0.449)	-0.506	(-1.671, 0.614)	-0.611	(-1.555, 0.235)
β_3^{π}	-0.302	(-1.419, 0.533)	-0.292	(-1.407, 0.578)	-0.265	(-1.314, 0.505)	-0.355	(-1.209, 0.372)
σ_1	1.535	(0.312, 4.338)	1.632	(0.381, 4.085)	1.517	(0.368, 4.542)	1.311	(0.065, 3.968)
σ_2	0.883	(0.045, 3.188)	0.831	(0.048, 2.777)	0.924	(0.037, 3.221)	0.646	(0.01, 2.516)
σ_3	0.902	(0.07, 2.804)	0.998	(0.152, 3.23)	0.883	(0.092, 3.001)	0.674	(0.024, 2.829)
$\pi_{global,1}$	0.255	(0.094, 0.585)	0.270	(0.102, 0.612)	0.247	(0.095, 0.541)	0.213	(0.081, 0.551)
$\pi_{global,2}$	0.376	(0.191, 0.569)	0.381	(0.193, 0.61)	0.382	(0.158, 0.649)	0.357	(0.174, 0.559)
$\pi_{global,3}$	0.430	(0.195, 0.63)	0.432	(0.197, 0.641)	0.438	(0.212, 0.624)	0.415	(0.23, 0.592)
$\pi_{os,1}$	0.214	(0.153, 0.276)	0.239	(0.181, 0.304)	0.220	(0.159, 0.285)	0.127	(0.044, 0.206)
$\pi_{os,2}$	0.423	(0.354, 0.493)	0.438	(0.369, 0.498)	0.428	(0.365, 0.492)	0.362	(0.291, 0.432)
$\pi_{os,3}$	0.508	(0.445, 0.568)	0.523	(0.456, 0.589)	0.510	(0.443, 0.572)	0.451	(0.381, 0.531)
$\pi_{pfs,1}$	0.095	(0.062, 0.142)	0.091	(0.061, 0.13)	0.094	(0.061, 0.135)	0.096	(0.06, 0.139)
$\pi_{pfs,2}$	0.343	(0.282, 0.402)	0.347	(0.292, 0.413)	0.344	(0.287, 0.406)	0.339	(0.28, 0.398)
$\pi_{pfs,3}$	0.412	(0.35, 0.47)	0.412	(0.354, 0.472)	0.414	(0.359, 0.47)	0.410	(0.354, 0.472)

1OS distribution/PFS distribution.

TABLE B6 Posterior summary statistics for hierarchical model parameters and all OS distributions with log-Normal PFS distribution.

	Exp/log-Normal[1]		Weibull/Log-Normal		Gompertz/Log-Normal		Log-logistic/Log-Normal	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
$\beta_{os,0}$	-3.085	(-3.248, -2.943)	3.027	(2.893, 3.19)	-3.086	(-3.244, -2.956)	2.802	(2.646, 2.985)
$\beta_{os,1}$	-0.014	(-0.023, -0.005)	0.011	(0.003, 0.02)	-0.013	(-0.022, -0.005)	0.010	(0, 0.019)
$\beta_{pfs,0}$	1.250	(1.206, 1.297)	1.249	(1.204, 1.301)	1.251	(1.205, 1.299)	1.252	(1.205, 1.299)
$\beta_{pfs,1}$	0.000	(-0.003, 0.003)	0.000	(-0.003, 0.003)	0.000	(-0.004, 0.004)	0.000	(-0.003, 0.003)
β_1^{π}	-1.134	(-1.973, 0.179)	-1.133	(-2.027, 0.034)	-1.175	(-2.171, 0.071)	-1.383	(-2.181, 0.117)
β_2^{π}	-0.379	(-1.346, 0.166)	-0.338	(-1.213, 0.314)	-0.366	(-1.339, 0.281)	-0.457	(-1.326, 0.208)
β_3^{π}	-0.109	(-1.035, 0.607)	-0.112	(-1.109, 0.525)	-0.104	(-1.075, 0.625)	-0.179	(-1.153, 0.438)
σ_1	1.114	(0.131, 3.274)	1.167	(0.17, 3.696)	1.141	(0.095, 3.668)	0.885	(0.017, 3.264)
σ_2	0.603	(0.013, 2.647)	0.570	(0.015, 2.431)	0.548	(0.011, 2.397)	0.586	(0.011, 2.643)
σ_3	0.629	(0.015, 2.941)	0.677	(0.034, 3.101)	0.569	(0.016, 2.282)	0.571	(0.008, 2.638)
$\pi_{global,1}$	0.256	(0.122, 0.545)	0.256	(0.116, 0.508)	0.250	(0.102, 0.518)	0.215	(0.101, 0.529)
$\pi_{global,2}$	0.410	(0.206, 0.541)	0.419	(0.229, 0.578)	0.413	(0.208, 0.57)	0.391	(0.21, 0.552)
$\pi_{global,3}$	0.475	(0.262, 0.647)	0.474	(0.248, 0.628)	0.476	(0.254, 0.651)	0.457	(0.24, 0.608)
$\pi_{os,1}$	0.212	(0.15, 0.273)	0.236	(0.172, 0.296)	0.216	(0.156, 0.282)	0.145	(0.069, 0.217)
$\pi_{os,2}$	0.433	(0.375, 0.489)	0.440	(0.381, 0.502)	0.429	(0.367, 0.487)	0.382	(0.3, 0.456)
$\pi_{os,3}$	0.509	(0.449, 0.567)	0.521	(0.458, 0.583)	0.509	(0.451, 0.565)	0.468	(0.396, 0.539)
$\pi_{pfs,1}$	0.143	(0.101, 0.189)	0.144	(0.098, 0.188)	0.145	(0.105, 0.19)	0.141	(0.103, 0.184)
$\pi_{pfs,2}$	0.420	(0.364, 0.469)	0.423	(0.37, 0.478)	0.420	(0.361, 0.479)	0.414	(0.356, 0.472)
$\pi_{pfs,3}$	0.487	(0.426, 0.545)	0.486	(0.43, 0.549)	0.485	(0.427, 0.543)	0.478	(0.419, 0.538)

1OS distribution/PFS distribution.

TABLE B7 Posterior summary statistics for separate model parameters and all OS distributions with exponential PFS distribution.

	Exp/Exp[1]		Weibull/Exp		Gompertz/Exp		Log-logistic/Exp	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
β_0^{os}	-3.042	(-3.203, -2.909)	3.002	(2.886, 3.141)	-3.051	(-3.197, -2.914)	2.715	(2.583, 2.858)
β_1^{os}	-0.012	(-0.021, -0.003)	0.009	(0.001, 0.018)	-0.012	(-0.021, -0.003)	0.007	(-0.001, 0.016)
β_0^{pfs}	-1.817	(-1.915, -1.719)	-1.815	(-1.917, -1.728)	-1.816	(-1.929, -1.712)	-1.817	(-1.919, -1.718)
β_1^{pfs}	-0.003	(-0.009, 0.004)	-0.002	(-0.009, 0.004)	-0.002	(-0.008, 0.005)	-0.003	(-0.009, 0.003)
$\beta_{os,1}^{\pi}$	-1.084	(-1.399, -0.816)	-1.008	(-1.32, -0.756)	-1.077	(-1.39, -0.787)	-1.308	(-1.699, -0.999)
$\beta_{os,2}^{\pi}$	-0.245	(-0.488, 0.014)	-0.190	(-0.429, 0.08)	-0.238	(-0.486, 0.017)	-0.398	(-0.715, -0.15)
$\beta_{os,3}^{\pi}$	0.068	(-0.185, 0.318)	0.102	(-0.125, 0.344)	0.070	(-0.183, 0.324)	-0.066	(-0.382, 0.196)
$\beta_{pfs,1}^{\pi}$	-1.848	(-2.174, -1.508)	-1.839	(-2.158, -1.542)	-1.829	(-2.171, -1.502)	-1.831	(-2.191, -1.507)
$\beta_{pfs,2}^{\pi}$	-0.626	(-0.878, -0.385)	-0.625	(-0.86, -0.398)	-0.617	(-0.881, -0.36)	-0.622	(-0.872, -0.372)
$\beta_{pfs,3}^{\pi}$	-0.361	(-0.595, -0.088)	-0.353	(-0.582, -0.11)	-0.348	(-0.545, -0.11)	-0.359	(-0.59, -0.117)
π_1^{os}	0.254	(0.198, 0.307)	0.268	(0.211, 0.32)	0.255	(0.199, 0.313)	0.214	(0.155, 0.269)
π_2^{os}	0.439	(0.38, 0.504)	0.453	(0.394, 0.52)	0.441	(0.381, 0.504)	0.402	(0.328, 0.463)
π_3^{os}	0.517	(0.454, 0.579)	0.525	(0.469, 0.585)	0.518	(0.454, 0.58)	0.484	(0.406, 0.549)
π_1^{pfs}	0.137	(0.102, 0.181)	0.138	(0.104, 0.176)	0.140	(0.102, 0.182)	0.139	(0.101, 0.181)
π_2^{pfs}	0.349	(0.294, 0.405)	0.349	(0.297, 0.402)	0.351	(0.293, 0.411)	0.350	(0.295, 0.408)
π_3^{pfs}	0.411	(0.355, 0.478)	0.413	(0.358, 0.472)	0.414	(0.367, 0.473)	0.411	(0.357, 0.471)

1OS distribution/PFS distribution.

TABLE B8 Posterior summary statistics for separate model parameters and all OS distributions with weibull PFS distribution.

	Exp/Weibull[1]		Weibull/Weibull		Gompertz/Weibull		Log-logistic/Weibull	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
β_0^{os}	-3.040	(-3.18, -2.913)	2.996	(2.877, 3.125)	-3.052	(-3.182, -2.914)	2.717	(2.585, 2.853)
β_1^{os}	-0.012	(-0.022, -0.003)	0.009	(0.001, 0.018)	-0.012	(-0.022, -0.003)	0.008	(0, 0.016)
β_0^{pfs}	1.780	(1.682, 1.887)	1.778	(1.683, 1.864)	1.774	(1.675, 1.869)	1.775	(1.683, 1.87)
β_1^{pfs}	0.001	(-0.005, 0.007)	0.000	(-0.006, 0.006)	0.001	(-0.005, 0.007)	0.000	(-0.004, 0.006)
$\beta_{os,1}^{\pi}$	-1.075	(-1.395, -0.786)	-1.004	(-1.315, -0.736)	-1.068	(-1.364, -0.792)	-1.309	(-1.659, -0.978)
$\beta_{os,2}^{\pi}$	-0.239	(-0.475, 0.005)	-0.188	(-0.43, 0.038)	-0.244	(-0.495, 0.004)	-0.401	(-0.711, -0.101)
$\beta_{os,3}^{\pi}$	0.062	(-0.196, 0.335)	0.107	(-0.18, 0.34)	0.071	(-0.184, 0.343)	-0.069	(-0.325, 0.218)
$\beta_{pfs,1}^{\pi}$	-1.772	(-2.083, -1.467)	-1.763	(-2.077, -1.437)	-1.770	(-2.116, -1.473)	-1.762	(-2.084, -1.432)
$\beta_{pfs,2}^{\pi}$	-0.562	(-0.813, -0.31)	-0.545	(-0.806, -0.301)	-0.543	(-0.791, -0.301)	-0.558	(-0.799, -0.312)
$\beta_{pfs,3}^{\pi}$	-0.323	(-0.562, -0.081)	-0.319	(-0.541, -0.115)	-0.319	(-0.557, -0.077)	-0.323	(-0.586, -0.102)
π_1^{os}	0.256	(0.199, 0.313)	0.269	(0.212, 0.324)	0.257	(0.204, 0.312)	0.214	(0.16, 0.273)
π_2^{os}	0.441	(0.383, 0.501)	0.453	(0.394, 0.51)	0.439	(0.379, 0.501)	0.402	(0.329, 0.475)
π_3^{os}	0.515	(0.451, 0.583)	0.527	(0.455, 0.584)	0.518	(0.454, 0.585)	0.483	(0.419, 0.554)
π_1^{pfs}	0.146	(0.111, 0.187)	0.148	(0.111, 0.192)	0.147	(0.108, 0.187)	0.148	(0.111, 0.193)
π_2^{pfs}	0.364	(0.307, 0.423)	0.368	(0.309, 0.425)	0.368	(0.312, 0.425)	0.365	(0.31, 0.423)
π_3^{pfs}	0.420	(0.363, 0.48)	0.421	(0.368, 0.471)	0.421	(0.364, 0.481)	0.420	(0.357, 0.474)

1OS distribution/PFS distribution.

TABLE B9 Posterior summary statistics for separate model parameters and all OS distributions with gompertz PFS distribution.

	Exp/Gompertz[1]		Gompertz		Gompertz/Gompertz		Log-logistic/Gompertz	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
β_0^{os}	-3.039	(-3.171, -2.896)	2.998	(2.887, 3.118)	-3.043	(-3.196, -2.915)	2.716	(2.582, 2.854)
β_1^{os}	-0.012	(-0.02, -0.003)	0.009	(0.001, 0.018)	-0.011	(-0.02, -0.001)	0.007	(-0.002, 0.016)
β_0^{pfs}	-1.815	(-1.903, -1.707)	-1.818	(-1.915, -1.721)	-1.814	(-1.914, -1.713)	-1.815	(-1.914, -1.725)
β_1^{pfs}	-0.003	(-0.009, 0.004)	-0.003	(-0.009, 0.004)	-0.002	(-0.009, 0.003)	-0.002	(-0.008, 0.004)
$\beta_{os,1}^{\pi}$	-1.072	(-1.344, -0.767)	-1.004	(-1.337, -0.718)	-1.076	(-1.368, -0.793)	-1.295	(-1.652, -0.976)
$\beta_{os,2}^{\pi}$	-0.242	(-0.479, -0.009)	-0.183	(-0.432, 0.051)	-0.221	(-0.434, -0.001)	-0.382	(-0.667, -0.116)
$\beta_{os,3}^{\pi}$	0.069	(-0.204, 0.34)	0.113	(-0.119, 0.388)	0.066	(-0.165, 0.294)	-0.067	(-0.326, 0.18)
$\beta_{pfs,1}^{\pi}$	-1.845	(-2.166, -1.564)	-1.834	(-2.156, -1.516)	-1.829	(-2.136, -1.519)	-1.822	(-2.134, -1.515)
$\beta_{pfs,2}^{\pi}$	-0.623	(-0.857, -0.383)	-0.624	(-0.874, -0.372)	-0.619	(-0.848, -0.403)	-0.624	(-0.856, -0.39)
$\beta_{pfs,3}^{\pi}$	-0.370	(-0.602, -0.125)	-0.362	(-0.583, -0.131)	-0.358	(-0.592, -0.124)	-0.347	(-0.58, -0.119)
π_1^{os}	0.256	(0.207, 0.317)	0.269	(0.208, 0.328)	0.255	(0.203, 0.311)	0.216	(0.161, 0.274)
π_2^{os}	0.440	(0.383, 0.498)	0.454	(0.394, 0.513)	0.445	(0.393, 0.5)	0.406	(0.339, 0.471)
π_3^{os}	0.517	(0.449, 0.584)	0.528	(0.47, 0.596)	0.516	(0.459, 0.573)	0.483	(0.419, 0.545)
π_1^{pfs}	0.138	(0.103, 0.173)	0.139	(0.104, 0.18)	0.139	(0.106, 0.18)	0.140	(0.106, 0.18)
π_2^{pfs}	0.350	(0.298, 0.405)	0.349	(0.294, 0.408)	0.351	(0.3, 0.401)	0.349	(0.298, 0.404)
π_3^{pfs}	0.409	(0.354, 0.469)	0.411	(0.358, 0.467)	0.412	(0.356, 0.469)	0.414	(0.359, 0.47)

1OS distribution/PFS distribution.

TABLE B10 Posterior summary statistics for separate model parameters and all OS distributions with log-logistic PFS distribution.

	Exp/Log-logistic[1]		Weibull/Log-logistic		Gompertz/Log-logistic		Log-logistic/Log-logistic	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
β_0^{os}	-3.047	(-3.201, -2.914)	2.995	(2.875, 3.127)	-3.048	(-3.183, -2.91)	2.717	(2.578, 2.867)
β_1^{os}	-0.012	(-0.021, -0.003)	0.009	(0.001, 0.018)	-0.012	(-0.02, -0.002)	0.008	(-0.001, 0.017)
β_0^{pfs}	1.328	(1.262, 1.395)	1.329	(1.26, 1.4)	1.331	(1.263, 1.406)	1.328	(1.253, 1.417)
β_1^{pfs}	0.002	(-0.002, 0.007)	0.002	(-0.002, 0.007)	0.002	(-0.003, 0.007)	0.002	(-0.002, 0.007)
$\beta_{os,1}^{\pi}$	-1.087	(-1.409, -0.804)	-1.001	(-1.3, -0.704)	-1.061	(-1.364, -0.799)	-1.300	(-1.69, -0.952)
$\beta_{os,2}^{\pi}$	-0.239	(-0.473, -0.007)	-0.200	(-0.423, 0.051)	-0.243	(-0.471, 0.006)	-0.383	(-0.661, -0.099)
$\beta_{os,3}^{\pi}$	0.060	(-0.169, 0.294)	0.092	(-0.13, 0.329)	0.069	(-0.17, 0.318)	-0.073	(-0.305, 0.195)
$\beta_{pfs,1}^{\pi}$	-1.870	(-2.199, -1.517)	-1.866	(-2.228, -1.535)	-1.862	(-2.201, -1.539)	-1.871	(-2.207, -1.566)
$\beta_{pfs,2}^{\pi}$	-0.618	(-0.875, -0.372)	-0.619	(-0.855, -0.372)	-0.621	(-0.858, -0.369)	-0.618	(-0.858, -0.356)
$\beta_{pfs,3}^{\pi}$	-0.360	(-0.596, -0.148)	-0.347	(-0.569, -0.12)	-0.348	(-0.568, -0.083)	-0.358	(-0.581, -0.122)
π_1^{os}	0.253	(0.196, 0.309)	0.270	(0.214, 0.331)	0.258	(0.204, 0.31)	0.216	(0.156, 0.279)
π_2^{os}	0.441	(0.384, 0.498)	0.450	(0.396, 0.513)	0.440	(0.384, 0.502)	0.406	(0.341, 0.475)
π_3^{os}	0.515	(0.458, 0.573)	0.523	(0.468, 0.581)	0.517	(0.458, 0.579)	0.482	(0.424, 0.549)
π_1^{pfs}	0.135	(0.1, 0.18)	0.135	(0.097, 0.177)	0.136	(0.1, 0.177)	0.135	(0.099, 0.173)
π_2^{pfs}	0.351	(0.294, 0.408)	0.351	(0.298, 0.408)	0.350	(0.298, 0.409)	0.351	(0.298, 0.412)
π_3^{pfs}	0.411	(0.355, 0.463)	0.414	(0.361, 0.47)	0.414	(0.362, 0.479)	0.412	(0.359, 0.47)

1OS distribution/PFS distribution.

TABLE B11 Posterior summary statistics for separate model parameters and all OS distributions with log-Normal PFS distribution.

	Exp/log-Normal[1]		Weibull/log-Normal		Gompertz/log-Normal		Log-logistic/log-Normal	
Parameter	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**	**Mean**	**95% CrI**
β_0^{os}	-3.046	(-3.183, -2.909)	2.996	(2.886, 3.113)	-3.039	(-3.186, -2.891)	2.714	(2.572, 2.868)
β_1^{os}	-0.012	(-0.02, -0.003)	0.009	(0.001, 0.017)	-0.012	(-0.02, -0.002)	0.007	(-0.002, 0.016)
β_0^{pfs}	1.247	(1.198, 1.294)	1.248	(1.197, 1.29)	1.249	(1.202, 1.295)	1.248	(1.204, 1.293)
β_1^{pfs}	0.000	(-0.004, 0.003)	0.000	(-0.003, 0.003)	0.000	(-0.003, 0.003)	0.000	(-0.003, 0.003)
$\beta_{os,1}^{\pi}$	-1.082	(-1.378, -0.824)	-1.002	(-1.346, -0.756)	-1.073	(-1.383, -0.775)	-1.292	(-1.71, -0.936)
$\beta_{os,2}^{\pi}$	-0.244	(-0.487, 0.022)	-0.176	(-0.423, 0.086)	-0.224	(-0.462, 0.011)	-0.398	(-0.7, -0.097)
$\beta_{os,3}^{\pi}$	0.056	(-0.178, 0.28)	0.110	(-0.107, 0.362)	0.062	(-0.192, 0.33)	-0.073	(-0.324, 0.193)
$\beta_{pfs,1}^{\pi}$	-1.575	(-1.904, -1.281)	-1.558	(-1.861, -1.298)	-1.565	(-1.889, -1.249)	-1.553	(-1.839, -1.268)
$\beta_{pfs,2}^{\pi}$	-0.307	(-0.515, -0.048)	-0.305	(-0.553, -0.066)	-0.307	(-0.519, -0.098)	-0.301	(-0.54, -0.056)
$\beta_{pfs,3}^{\pi}$	-0.067	(-0.321, 0.215)	-0.071	(-0.303, 0.189)	-0.060	(-0.309, 0.176)	-0.056	(-0.273, 0.164)
π_1^{os}	0.254	(0.201, 0.305)	0.270	(0.207, 0.32)	0.256	(0.201, 0.315)	0.217	(0.153, 0.282)
π_2^{os}	0.440	(0.381, 0.506)	0.456	(0.396, 0.521)	0.444	(0.387, 0.503)	0.402	(0.332, 0.476)
π_3^{os}	0.514	(0.456, 0.57)	0.527	(0.473, 0.59)	0.515	(0.452, 0.582)	0.482	(0.42, 0.548)
π_1^{pfs}	0.173	(0.13, 0.217)	0.175	(0.135, 0.215)	0.174	(0.131, 0.223)	0.176	(0.137, 0.22)
π_2^{pfs}	0.424	(0.374, 0.488)	0.425	(0.365, 0.484)	0.424	(0.373, 0.476)	0.426	(0.368, 0.486)
π_3^{pfs}	0.483	(0.42, 0.554)	0.482	(0.425, 0.547)	0.485	(0.423, 0.544)	0.486	(0.432, 0.541)

1OS distribution/PFS distribution.

TABLE B12 WAIC statistics for all distributions and hierarchical model.

		ELPD†		$p_D^{\ddagger}$		WAIC	
OS	**PFS**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
exp	exp	-5064.70	80.37	11.13	0.46	10129.40	160.74
exp	gompertz	-5065.64	80.44	11.80	0.49	10131.29	160.88
exp	loglogistic	-4970.34	80.72	11.78	0.33	9940.68	161.45
exp	lognormal	-4730.47	92.79	18.04	0.91	9460.94	185.59
exp	weibull	-5062.99	81.69	14.63	0.80	10125.97	163.38
gompertz	exp	-5064.99	80.41	11.45	0.45	10129.99	160.82
gompertz	gompertz	-5065.61	80.44	11.66	0.48	10131.22	160.88
gompertz	loglogistic	-4970.42	80.79	11.68	0.34	9940.84	161.59
gompertz	lognormal	-4730.92	92.64	18.31	0.95	9461.84	185.29
gompertz	weibull	-5063.81	81.74	15.45	0.85	10127.61	163.48
loglogistic	exp	-5059.86	80.47	11.29	0.42	10119.73	160.93
loglogistic	gompertz	-5060.37	80.40	11.64	0.45	10120.74	160.80
loglogistic	loglogistic	-4966.22	80.75	12.77	0.35	9932.44	161.50
loglogistic	lognormal	-4726.57	93.17	19.43	0.91	9453.13	186.34
loglogistic	weibull	-5058.17	81.84	14.80	0.74	10116.33	163.67
lognormal	exp	-4972.55	84.58	16.35	0.61	9945.09	169.15
lognormal	gompertz	-4973.16	84.69	16.50	0.56	9946.32	169.38
lognormal	loglogistic	-4877.87	85.79	16.53	0.50	9755.74	171.59
lognormal	lognormal	-4638.85	100.27	23.58	1.00	9277.71	200.55
lognormal	weibull	-4970.22	86.17	19.01	0.89	9940.44	172.34
weibull	exp	-5063.26	80.70	11.93	0.46	10126.53	161.40
weibull	gompertz	-5064.37	80.64	12.67	0.52	10128.73	161.28
weibull	loglogistic	-4968.74	81.03	12.30	0.37	9937.49	162.07
weibull	lognormal	-4729.78	93.18	19.54	0.97	9459.56	186.36
weibull	weibull	-5061.91	82.05	15.77	0.85	10123.81	164.11

†ELPD: Expected log pointwise predictive density; ‡p_D: Effective number of parameters

TABLE B13 WAIC statistics for all distributions and separate model.

		ELPD†		$p_D^{\ddagger}$		WAIC	
OS	**PFS**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
exp	exp	-5066.38	81.40	10.46	0.43	10132.76	162.81
exp	gompertz	-5066.68	81.46	10.38	0.46	10133.37	162.92
exp	loglogistic	-4971.73	81.67	10.36	0.28	9943.46	163.34
exp	lognormal	-4732.41	93.18	18.56	0.99	9464.83	186.36
exp	weibull	-5064.87	82.86	14.39	0.86	10129.74	165.72
gompertz	exp	-5066.72	81.44	10.51	0.45	10133.43	162.89
gompertz	gompertz	-5066.76	81.49	10.19	0.44	10133.53	162.97
gompertz	loglogistic	-4972.24	81.70	10.70	0.28	9944.47	163.40
gompertz	lognormal	-4732.23	93.30	18.01	0.95	9464.46	186.61
gompertz	weibull	-5064.46	82.92	13.50	0.78	10128.91	165.83
loglogistic	exp	-5062.51	81.89	10.79	0.40	10125.02	163.78
loglogistic	gompertz	-5062.64	82.01	10.26	0.41	10125.27	164.02
loglogistic	loglogistic	-4968.45	82.36	11.63	0.28	9936.89	164.72
loglogistic	lognormal	-4728.25	94.20	18.18	0.92	9456.51	188.40
loglogistic	weibull	-5060.18	83.31	13.79	0.69	10120.37	166.62
lognormal	exp	-4973.43	85.26	14.63	0.55	9946.86	170.53
lognormal	gompertz	-4973.48	85.22	14.31	0.51	9946.96	170.44
lognormal	loglogistic	-4879.07	86.55	15.07	0.44	9758.14	173.09
lognormal	lognormal	-4638.43	100.40	21.30	0.96	9276.85	200.80
lognormal	weibull	-4972.48	87.07	19.54	0.91	9944.95	174.13
weibull	exp	-5064.74	81.52	11.11	0.45	10129.48	163.04
weibull	gompertz	-5065.16	81.59	11.16	0.44	10130.32	163.18
weibull	loglogistic	-4970.64	81.87	11.68	0.35	9941.27	163.73
weibull	lognormal	-4730.93	93.78	19.04	0.96	9461.85	187.56
weibull	weibull	-5063.20	83.09	14.87	0.83	10126.41	166.17

†ELPD: Expected log pointwise predictive density; ‡p_D: Effective number of parameters

TABLE B14 PSIS-LOO statistics for all distributions and hierarchical model.

		ELPD[†]		$p_D^{\ddagger}$		LOO AIC	
OS	**PFS**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
exp	exp	-5064.80	80.37	11.22	0.46	10129.59	160.75
exp	gompertz	-5065.74	80.44	11.90	0.49	10131.48	160.88
exp	loglogistic	-4970.43	80.73	11.87	0.33	9940.86	161.45
exp	lognormal	-4730.60	92.80	18.18	0.92	9461.20	185.59
exp	weibull	-5063.08	81.69	14.72	0.80	10126.17	163.38
gompertz	exp	-5065.09	80.41	11.55	0.45	10130.18	160.82
gompertz	gompertz	-5065.70	80.44	11.75	0.48	10131.40	160.88
gompertz	loglogistic	-4970.50	80.79	11.76	0.34	9940.99	161.59
gompertz	lognormal	-4731.06	92.65	18.45	0.96	9462.12	185.30
gompertz	weibull	-5063.92	81.74	15.57	0.86	10127.84	163.49
loglogistic	exp	-5059.95	80.47	11.38	0.42	10119.90	160.94
loglogistic	gompertz	-5060.45	80.40	11.73	0.45	10120.90	160.80
loglogistic	loglogistic	-4966.31	80.75	12.87	0.35	9932.63	161.50
loglogistic	lognormal	-4726.71	93.18	19.58	0.92	9453.42	186.35
loglogistic	weibull	-5058.26	81.84	14.90	0.74	10116.53	163.67
lognormal	exp	-4972.68	84.58	16.49	0.61	9945.37	169.16
lognormal	gompertz	-4973.28	84.69	16.62	0.57	9946.56	169.39
lognormal	loglogistic	-4878.00	85.79	16.67	0.50	9756.01	171.59
lognormal	lognormal	-4639.02	100.27	23.75	1.00	9278.04	200.55
lognormal	weibull	-4970.34	86.17	19.13	0.89	9940.68	172.34
weibull	exp	-5063.35	80.70	12.02	0.47	10126.69	161.40
weibull	gompertz	-5064.46	80.64	12.76	0.53	10128.92	161.28
weibull	loglogistic	-4968.83	81.03	12.39	0.37	9937.67	162.07
weibull	lognormal	-4729.90	93.18	19.66	0.98	9459.80	186.36
weibull	weibull	-5062.01	82.05	15.88	0.85	10124.03	164.11

[†]ELPD: Expected log pointwise predictive density; [‡]p_D: Effective number of parameters

TABLE B15 PSIS-LOO statistics for all distributions and separate model.

		ELPD[†]		$p_D^{\ddagger}$		LOO AIC	
OS	**PFS**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
exp	exp	-5066.47	81.40	10.55	0.43	10132.93	162.81
exp	gompertz	-5066.76	81.46	10.46	0.46	10133.52	162.92
exp	loglogistic	-4971.81	81.67	10.44	0.28	9943.61	163.34
exp	lognormal	-4732.55	93.18	18.69	0.99	9465.09	186.37
exp	weibull	-5064.95	82.86	14.48	0.86	10129.91	165.72
gompertz	exp	-5066.80	81.44	10.59	0.46	10133.60	162.89
gompertz	gompertz	-5066.84	81.49	10.27	0.44	10133.69	162.97
gompertz	loglogistic	-4972.33	81.70	10.79	0.29	9944.66	163.40
gompertz	lognormal	-4732.34	93.31	18.13	0.95	9464.68	186.61
gompertz	weibull	-5064.54	82.92	13.58	0.79	10129.08	165.84
loglogistic	exp	-5062.60	81.89	10.88	0.40	10125.20	163.78
loglogistic	gompertz	-5062.72	82.01	10.34	0.41	10125.43	164.03
loglogistic	loglogistic	-4968.53	82.36	11.72	0.29	9937.06	164.72
loglogistic	lognormal	-4728.37	94.20	18.30	0.93	9456.75	188.40
loglogistic	weibull	-5060.29	83.31	13.90	0.69	10120.59	166.62
lognormal	exp	-4973.54	85.26	14.74	0.56	9947.08	170.53
lognormal	gompertz	-4973.58	85.22	14.40	0.52	9947.15	170.44
lognormal	loglogistic	-4879.18	86.55	15.17	0.45	9758.35	173.10
lognormal	lognormal	-4638.57	100.40	21.44	0.96	9277.15	200.80
lognormal	weibull	-4972.61	87.07	19.67	0.92	9945.22	174.14
weibull	exp	-5064.82	81.52	11.18	0.45	10129.63	163.04
weibull	gompertz	-5065.24	81.59	11.24	0.44	10130.48	163.18
weibull	loglogistic	-4970.70	81.87	11.74	0.35	9941.40	163.73
weibull	lognormal	-4731.06	93.78	19.18	0.97	9462.12	187.57
weibull	weibull	-5063.28	83.09	14.95	0.82	10126.56	166.18

[†]ELPD: Expected log pointwise predictive density; [‡]p_D: Effective number of parameters