Katholieke Universiteit Leuven

Bayesian Data Analysis II

Project: Part 1

Modelling white grub worms' survival against entomopathogenic nematodes

Authors:
Hassler Valentin, r0876714
name@student.kuleuven.be

Benson Leo, r0875515
carlleo.benson@student.kuleuven.be

Dema Juli, r0821709
juli.dema@student.kuleuven.be

Fernandez Diego, r0817518
diego.fz@student.kuleuven.be

Bruno Mattia, r0874568
mattia.bruno@student.kuleuven.be

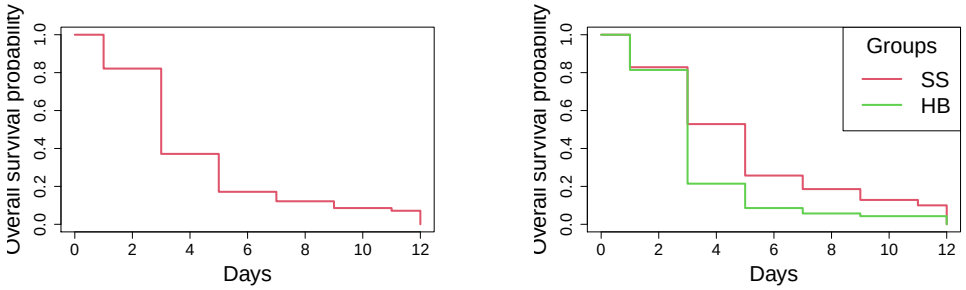
January 14, 2022

1 Introduction

Entomopathogenic nematodes (EPNs) are an alternative to chemical pesticides against the pests white grubs. To evaluate the effectiveness of the EPNs, two different species of EPNs were used in experiments: S. sacchari (SS) and H. baujardi (HB). The data provided for this analysis is collected as follows: 140 grubs were observed over a period of 12 days, whereby 7 grubs each were placed on each of the 20 different cell culture plates with a gel-soil mixture. EPNs of the species SS and HB were used for 10 of the cell culture plates each. At the beginning of the experiment, the size of all grubs was measured. The observations made during the 12 days, at intervals of 2 days each, consisted of checking whether the larvae were still alive or had already died. Thus, the dataset contains 5 covariates: group (1 for SS, 2 for HB), grupsize (cm), urepid (plate id, 1-20), lowerlim (lower bound of interval where grub died), upperlim (upper bound of interval where grub died, missing value denotes no limit). We are interested in finding out if a log-normal or a Weibull are suitable for the grub survival distribution and if the latter depends on the grubs' size and on the EPN species used. For modelling the effectiveness of the EPNs against the white grubs, survival analysis is appropriate, employing survival models. The computational tasks which are required from survival modelling are handily executed using packages in R, particularly rjags and coda.

Look at the median death time per EPN species to assess which of the two EPN species acts the fastest in killing the white grubs.

If we look at the data, we see that the median time of death for the grubs that have come into contact with the EPNs of the SS species is between the fourth and sixth day. At this point, only a good 20% of the grubs of species HB are still alive, as can also be seen in Figure 1 . The median time of death for the grubs of species HB is between the second and fourth day. Therefore, we already have a first indication that it is more likely that the EPNs of species HB are faster in killing the grubs. However, as this is only a rough summary statistic, we will have to examine this hypothesis more closely in the further parts of the paper.



(a) Total survival

(b) Survival per Group Curves

Figure 1: Survival Curves

The remaining questions are processed separately for two different forms of the variable to be explained (survival time). First, a simplified dataset is used in which the respective survival time is set to the mean value between the penultimate and last observation time before the death of the grub. Or, if the grub is still alive at the end, the time of death is set to the twelfth day. Second, the original data are considered (with interval and right censoring). This subdivision is carried out on the basis of the given task. For all the rest of the analysis, we use the size of the constrictors in normalised form, as this makes the results between the intercept and the beta parameter less correlated, and thus the model has less autocorrelation. Therefore, fewer iterations need to be performed for each simulated Posterior Chain to obtain reliable results.

2 Non-censored model

Evaluate the effect of the covariate grubs' size.

To evaluate the effect of the size of the grubs' size in the median survival time (T), the following model was proposed:

$$log(T_i) = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$$

where for each subject $i = (1, \ldots, N)$, $x_{i,2}$ represents the kind of EPN used and $x_{i,1}$, the covariate related to the grubs' size, which is normalized to accelerate convergence. Every subject i shares the same baseline value β_0 and distribution of its random error $\epsilon_i \sim N(0, \sigma^2)$.

This model was tested using jags, with a burn-in of size 2000 and a sampling size of 5000. There was no thinning in this case for neither of the three chains [1].We have chosen a normal distribution with a mean of zero and a variance of 10^6 as prior for the different beta factors in order to avoid prescribing to the MCMC algorithm in which range it should place the beta values. For the random errors, we assumed an inverse gamma distribution for the variance with shape and scale parameters of size 10^{-3}. We have chosen the given variance because it replicates Jeffrey's prior for the variance of a normal distribution, thereby ensuring the invariability of the prior information by transformation and still allowing all values between 0 and infinity to be taken. However, the sensitivity analysis carried out shows that the results are very robust and that even with a very different choice of prior, approximately the same results are still obtained.
The convergence of the model was assured with formal tests. The Gelman test resulted in the estimated potential scale reduction factor equal to 1, for all parameters. The Geweke dynamic test resulted in most Z-scores within range for both chains, as it can be seen in Figure 2 for the first chain.

It can be seen from Figure 3 that there is no clear effect of the size of the grub.

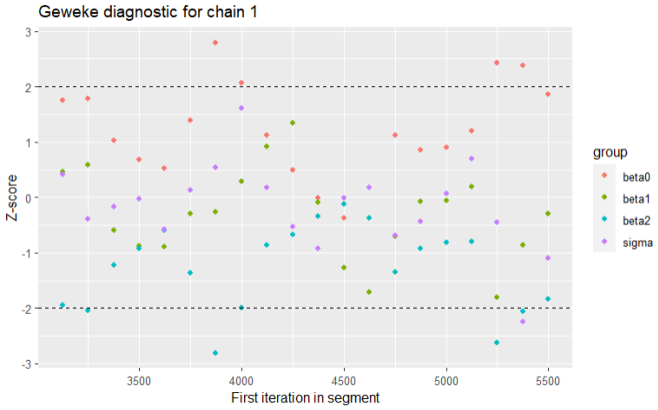


Figure 2: Geweke diagnostic for the first chain of the Lognormal model.

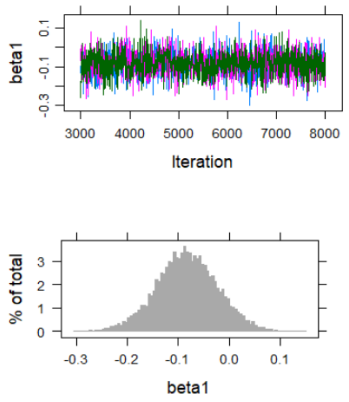


Figure 3: Traceplot and histogram of sampled values of β_1

This is also evident from the Table 1, where the 95% credibility interval clearly contains zero.

[1]All chains were initialised with different start values. After 2000 burn-in iterations, however, no difference could be seen between the patterns of the individual courses of the chains. This was also the case for all other models in this term paper.

	Mean	SD	2.5%	97.5%	MCSE
β_0	1.36	0.0844	1.191	1.524	0.00157
β_1	-0.0833	0.0602	-0.203	0.0351	0.000673
β_2	-0.31	0.12	-0.543	-0.0788	0.0022
σ	0.6964	0.0424	0.619	0.786	0.000354

Table 1: 95% credible interval for the model parameters

In addition to the lognormal model considered so far, a Weibull model is also considered. The model proposed in this context is specified as follows:

$$log(T_i) = \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$$

with $T_i \sim W(\gamma, \lambda_i)$ and ϵ_i following a type I least extreme value distribution centered at β_0 with scale parameter σ, where[2]:

$$\begin{aligned} \lambda_i &= (\alpha_i^{-1})^{\gamma} \\ \alpha_i^{-1} &= e^{-\mu_i} \\ \mu_i &= \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} \end{aligned}$$

The sampling settings for this case were similar as the previous case. The choice of prior distribution for the Weibull model was analogous to the choice of prior distribution for the lognormal model. When choosing the prior of the shape parameter, we decided to choose a uniform prior with min 0 and max 100. This choice is intended to symbolise no particular preference for a specific size of the shape parameter. The sensitivity analysis also shows that the results of the individual estimated parameters are robust. Again, the convergence was assured with formal tests. The Gelman test resulted in the estimated potential scale reduction factor equal to 1, for all parameters. The Geweke dynamic test resulted in most Z-scores within range for both chains, as it can be seen in Figure 4 for the first chain.

Figure 5 and Table 2 show that the model does not indicate a clear effect of the grubs' size on the expected survival time. This can be seen, since both the figure of the marginal posterior predictive distribution and the table show that the 95% credible interval of the parameter β_1 contain the value 0.
The DIC value of the lognormal model set up is 636.7, which is about 4.5 units lower than the value we observed for the Weibull model (641.1). Therefore, the observed

[2]The Weibull distribution is therby parameterized as in the case of JAGS, different from R

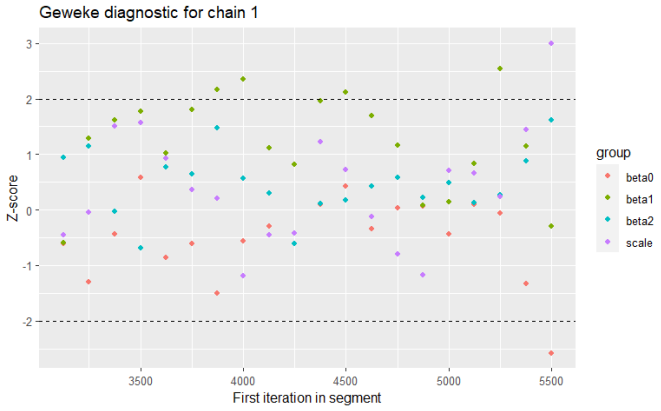


Figure 4: Geweke diagnostic for the first chain of the Weibull model.

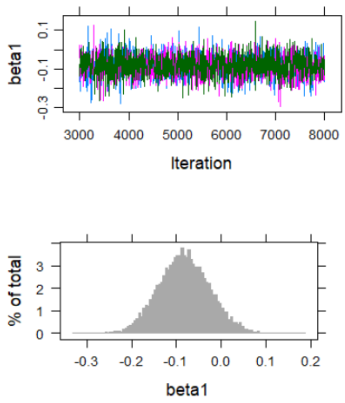


Figure 5: Traceplot and histogram of sampled values of β_1

	Mean	SD	2.5%	97.5%	MCSE
β_0	1.725	0.0768	1.575	1.874	0.00138
β_1	-0.0827	0.0575	-0.193	0.0323	0.000604
β_2	-0.374	0.106	-0.58	-0.167	0.00186
scale	0.631	0.0395	0.56	0.715	0.000448

Table 2: 95% credible interval for the model parameters

DIC values support favouring the lognormal model over the Weibull model. The pseudo Bayes factor is basically equivalent between the lognormal model (LPML value is -2.2726) over the Weibull model (LPML value is -2.2927), as the observed pseudo logarithmic marginal likelihood for the lognormal model is about 0.02 units larger than that for the observed Weibull model. This means that the pseudo Bayes factor takes on the approximate value of 1.02.

To see how realistically the models can be responsible for generating the observed data, we perform posterior predictive checks. We compare the observed data with artificial survival time data, which we generate using the posterior predictive distribution. However, since there is a double data use Bayarri and Berger 2000, the estimated values have to be treated with additional caution (for example, overfitting might not be detected). To perform a posterior predictive check, we compare the number of observed survivals above 12 days with the estimated number of grubs that would still be alive after 12 days given the same available data. We find that using the lognormal model, in about 7.9% of the cases at least as many grubs survive

12 days as is the case for the observed data (10 grubs). For the Weibull model simulations, it is about 0.7%. This observation therefore suggests that it is unlikely that the observed data were generated by the specified Weibull model. When we repeated this test for the grubs that survived shorter than 2 days, we observed that in the lognormal model it is estimated in 92.7% of the cases that at least as many grubs have already died as we observe in the data. In the Weibull model, it is about 92.5% of the cases. Therefore, the posterior predictive checks performed are in line with the observed DIC values, which also favour the lognormal model over the Weibull model. Figure 6 shows the estimates of survival times in the lognormal

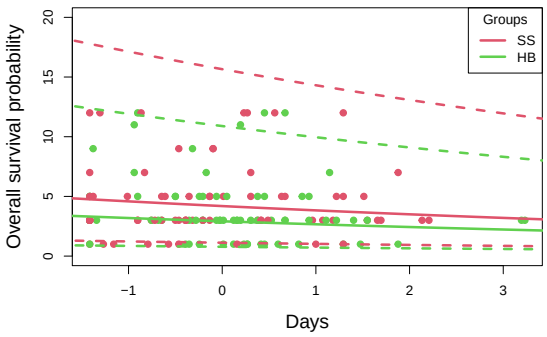


Figure 6: Estimated Survival Time in the Lognormal Model

model. It can be seen that the survival time for grubs in contact with EPNs of the SB species is estimated to be about one day higher than for the HB species. The wide 95% confidence intervals indicate that the estimates are relatively imprecise. [3]

Consider the clustering by including a random effect, i.e., the variation of location parameters across replicate plates. What does the random effect's variance component estimate suggest about the "clustering" of responses?

To introduce random effects into the noncensored model, we consider a lognormal AFT model, modelling survival time T with added random intercept b_{0i}:

$$log(T_{ij}) = \beta_0 + \beta_1 x_{ij,1} + \beta_2 x_{ij,2} + b_{0i} + \epsilon_{ij},\ i = 1,\ldots,20,\ j = 1,\ldots,7$$
$$\text{where } b_{0i} \sim N(0,\sigma_{b_0}),\ \epsilon_{ij} \sim N(0,\sigma^2)$$

[3]The estimates were constructed using the mean values of the sampled parameters, these were then plugged into the Theoretical Quantile Function to obtain the results presented

where x_{ij1} and x_{ij2} once again represent the covariates grub size and group, respectively, and i represents the unique plates (1-10 group 1, 11-20 group 2) and j the repeats. Priors used were

$$\sigma^2 = IG(10^{-3},10^{-3}),\ \beta_0,\beta_1,\beta_2 \sim N(0,10^3),\ \sigma_{b_0} \sim U(0,100)$$

Like for the previous models, this model was tested with a burn-in of 2000, sample-size of 5000 and 3 chains. The reason for these priors are motivated in the same way as the non-random effects model, with a vague prior for b_0 also being chosen.
A sensitivity analysis was conducted by varying priors for the different parameters (Table 3) which produced quite homogenous results across all settings for priors, suggesting that the posterior is more dependent on the likelihood than the priors in the model.

	Baseline	$\sigma^2_\beta = 10^3$	$\sigma^2_\beta = 10^9$	$\sigma^2 \sim IG(0.2,0.2)$	$\sigma^2 \sim IG(10^{-4},10^{-4})$ $\sigma^2_{b_0} \sim G(10^{-2},10^{-2})$	$\beta \sim N(0,10^9)$ $\sigma^2_{b_0} \sim G(10^{-2},10^{-2})$	$\beta \sim N(0,10^3)$ $\sigma \sim IG(10{-}3,10^{-3})$
σ	0.6(0.04)	0.6(0.04)	0.6(0.04)	0.59(0.04)	0.59(0.04)	0.6(0.04)	0.59(0.04)
β_0	1.35(0.16)	1.36(0.16)	1.35(0.15)	1.35(0.15)	1.34(0.14)	1.36(0.15)	1.35(0.16)
β_1	-0.05(0.05)	-0.05(0.05)	-0.05(0.05)	-0.05(0.05)	-0.05(0.05)	-0.05(0.05)	N.A.
β_2	-0.31(0.23)	-0.31(0.22)	-0.3(0.22)	-0.29(0.22)	-0.28(0.2)	-0.31(0.21)	-0.29(0.22)
$\sigma^2_{b_0}$	0.2(0.1)	0.2(0.11)	0.2(0.1)	0.2(0.1)	0.16(0.08)	0.16(0.08)	0.21(0.11)

Table 3: Sensitivity analysis for log-normal random effect model. Entries are given as "Mean(SD)". Columns indicate used priors. Baseline priors are assumed unless stated otherwise, these are: $\beta \sim N(0,10^3),\ \sigma^2 \sim U(0,1000)\ \sigma^2_{b_0} \sim U(0,100)$. The rightmost model also removed the covariate grub size (x_1).

The convergence of the model was checked by Geweke diagnostic, Gelman test, and traceplots. The Geweke diagnostic resulted in the large majority of parameters being within the 2 SD bounds with a few outliers, particularly for β_2, and somewhat for β_0 and σ^2, however convergence can still be assumed (Fig. 7). As seen in Table 4, we have fairly certain results for all parameters but β_0 and β_2 both have fairly high autocorrelation, but still the effective amount of samples remain at an acceptable size. We also see a clear difference in the autocorrelation between different groups, as group 1 (index 1 to 10) in general shows about twice the autocorrelation of group 2 (index 11 to 20). Moreover, β_1 is estimated to be quite small, with 0 within range of SD still (like for log-normal without random intercept) suggesting a weak link between the outcome and the covariate grub size. Since the magnitude of the between class variability σ^2_{b0} is only about half of the within one σ^2, we can conclude that different plates have a different effect on the survival time.

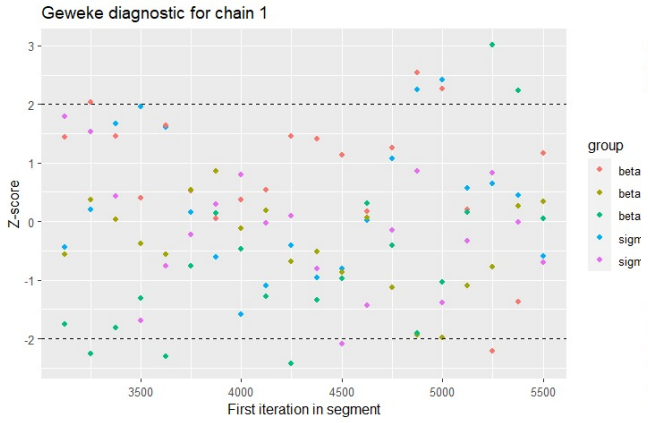


Figure 7: Geweke diagnostic for the first chain of the lognormal random effect model.

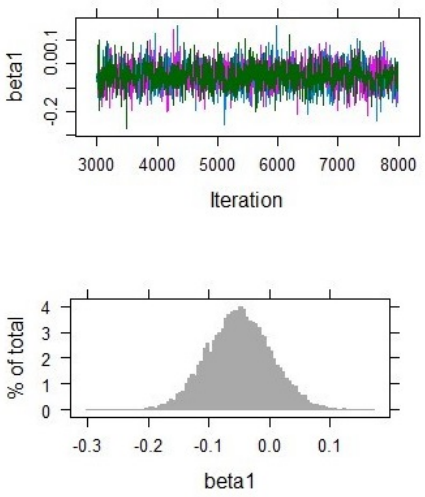


Figure 8: Traceplot and histogram of sampled values of β_1

	Lower95	Median	Upper95	Mean	SD	MCerr	MC%ofSD	SSeff	AC.10
σ^2	0.268	0.346	0.443	0.346	0.045	0	1	9682	-0.005
β_0	1.024	1.36	1.675	1.356	0.165	0.008	4.6	475	0.537
β_1	-0.154	-0.05	0.057	-0.051	0.054	0.001	1.2	7347	-0.002
β_2	-0.739	-0.309	0.155	-0.305	0.226	0.01	4.5	494	0.51
$\sigma^2_{b_0}$	0.056	0.181	0.407	0.203	0.104	0.002	2	2524	0.107

Table 4: Summary of posterior parameters for lognormal random-intercept model.

To further check the plausibility of the model, a posterior predictive checks are done on the model. Like before, the maximum, minimum, Kolmogorov-Smirnov, and Sinharay and Stern statistics were used, and neither of the statistics suggest a lack of fit (Table 5).
A Weibull AFT model with a random intercept is also considered:

$$log(T_{ij}) = \beta_1 x_{ij,1} + \beta_2 x_{ij,2} + b_{0i} + \epsilon_{ij},$$

with $T_{ij} \sim W(\gamma, \lambda_{ij})$, $b_{0i} \sim N(0,\sigma_{b_0})$ and ϵ_{ij} following a type I least extreme value

Statistic	p_D
Maximum	0.75
Minimum	0.76
Kolmogorov-Smirnov	0.33
Sinharay and Stern	0.83

Table 5: Posterior predictive p-values for various GOF statistics on log-normal random intercept model.

distribution centered at β_0 with scale parameter σ, where:

$$\begin{aligned} \lambda_{ij} &= \alpha_{ij}^{-1}\gamma \\ \alpha_{ij}^{-1} &= e^{-\mu_{ij}} \\ \mu_{ij} &= \beta_0 + \beta_1 x_{ij,1} + \beta_2 x_{ij,2} + b_{0i} \end{aligned}$$

where the same settings as the lognormal random effect model were used, along with priors

$$\sigma_{b_0}^2 \sim U(0{,}100),\ \gamma \sim U(0{,}100),\ \beta \sim N(0{,}10^6).$$

Posteriors for the model parameters are shown in Table 6. Similarly to the log-normal random intercept model we see that the covariate grub size seems to play a small role in determining the outcome since 0 is still included within the standard deviation of β_1. Apart from this, the covariate group seems to have less of an impact in this model as compared to the log-normal random intercept model, since β_2 nearly contains 0 within its standard deviation. Like the log-normal random intercept model, the random intercept variance has a significant effect, and thus effects between plates are likely to exist. The shape parameter γ is estimated to be quite high. One interpretation of this is that the hazard rate increases over time, so it is more likely that the grub dies the longer the experiment is conducted.

	Lower95	Median	Upper95	Mean	SD	MCerr	MC%ofSD	SSeff	AC.10
γ	1.57	1.84	2.09	1.84	0.13	0.00	1.40	4887	0.01
β_0	1.30	1.63	1.93	1.63	0.16	0.01	4.90	420	0.57
β_1	-0.16	-0.05	0.05	-0.06	0.05	0.00	1.20	6862	-0.00
β_2	-0.72	-0.27	0.21	-0.27	0.23	0.01	5.00	405	0.58
$\sigma_{b_0}^2$	0.05	0.21	0.48	0.24	0.12	0.00	2.00	2414	0.10

Table 6: Summary of posterior parameters for Weibull random-intercept model.

Further, convergence was shown with Geweke diagnostics (Fig. 9), Gelman test, and traceplots, like for previous models.

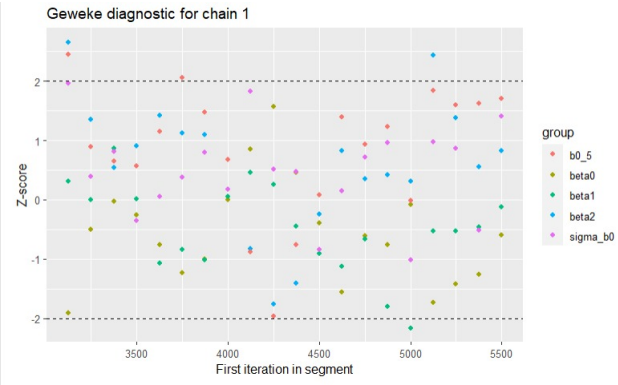


Figure 9: Geweke diagnostic for the first chain of the Weibull random effect model.

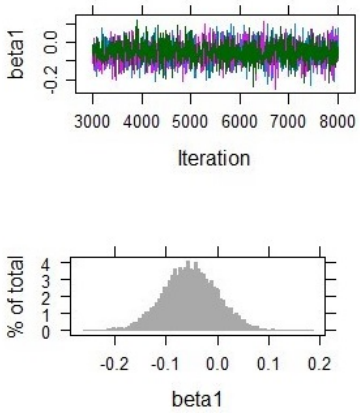


Figure 10: Traceplot and histogram of sampled values of β_1

Posterior predictive checks were also performed with the same statistics as before, with results once more suggesting no lack of fit (Table 7). As in the case without

Statistic	p_D
Maximum	0.76
Minimum	0.8
Kolmogorov-Smirnov	0.29
Sinharay and Stern	0.87

Table 7: Posterior predictive p-values for various GOF statistics on log-normal random intercept model.

random effect, the DIC criteria for the log-normal (605.3) is smaller than the Weibull one (619.2), showing that the first one better fits data. The pseudo Bayes factor has a value of 1.08 in this case favouring the Lognormal model by the slightest margin.

Check whether there are outlying/influential observation(s). If so, give possible reasons for the outlying observation(s)

For the log-normal model with random effects, outlying/influential observation(s) are detected using conditional predictive ordinates (CPO), which gives a cross-validatory

measure of extremeness. Figure 11 shows the plot of the inverse-CPO estimates. We can see that the observation 105 seems to represent an outlier and also observations 70, 91, 98, 126 have high inverse-CPO values. All these worms (except the 98) were originally right-censored, therefore taking 12 days as survival time for them makes the estimate unreliable. Therefore, a possible reason for these outlying observations is that especially long survival times cannot be modelled adequately since there are only small differences in the data used apparent for estimating survival times and the flexibility of the models set up is also limited. The Posterior Predict Ordinate values showed a similar pattern, though the outliers were somewhat less extreme in the case of the PPO.

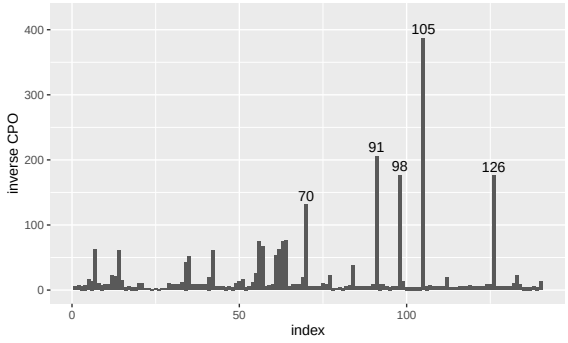


Figure 11: Index plot of inverse-CPO estimates

Check the distribution of the random effect and propose a robust distribution for the random effect.

Each random intercept is sampled repeatedly and a histogram of the posterior means of each one is shown below in Figure 12 (left). The distribution does not seem normal at a glance. This can occur for various reasons, for example due to an important covariate being omitted (Lesaffre and Lawson 2012). However, it is difficult to check the distributional assumption of normality through the histogram since the number of posterior means is small (only 20) and a clear picture cannot be obtained. Thus, a Q-Q plot which plots the quantiles of the data against the theoretical quantiles of a normal distribution is also shown in Figure 12 (right). The Q-Q plot shows a decent fit for most of the random effects but clear outliers are present.

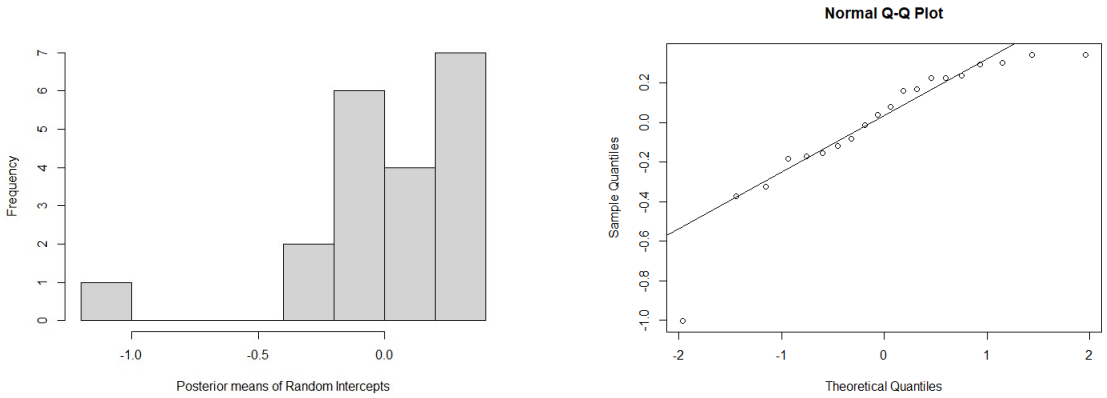


Figure 12: Histogram of the posterior means of the random intercepts (left) and Q-Q plot of the quantiles of the random intercept posterior means against the theoretical quantiles of a normal distribution (right).

Further, Posterior Predictive Checks (PPC) are performed to test the goodness of fit of the assumed normal distribution. These checks compare a statistic evaluated once on the data and once on the sampled posterior predictive distribution. A large discrepancy between the two indicates a poor fit. Here, the following statistics are considered: *minimum, maximum, Kolmogorov-Smirnov statistic* and *Sinharay and Stern statistic*. The results are summarized below, in Table 8. Since the PPP-values are larger than 0.05, none of them indicate a poor fit of the normal distribution to the random intercepts.

Statistic	p_D
Maximum	0.75
Minimum	0.77
Kolmogorov-Smirnov	0.33
Sinharay and Stern	0.84

Table 8: Posterior predictive p-values for various GOF statistics.

Nonetheless, an appropriate robust distribution could be a t-distribution with low degrees of freedom because outliers are downweighted (Lange, Little, and Taylor 1989). A t-distribution with 4 degrees of freedom is chosen based on the paper by Diya et al. (2012). The histogram of the posterior means of the random intercepts is shown in Figure 13 (left) and the respective Q-Q plot is shown in Figure 13 (right).

From the figures, an imporvement in the fit is not obvious. Further, the difference in the DIC when specifiying a t-distribution with 4 degrees of freedom for the random effects (DIC = 248.9) compared to a normal distribution (DIC = 250) is not great. This could be due to the outliers or influential observations.

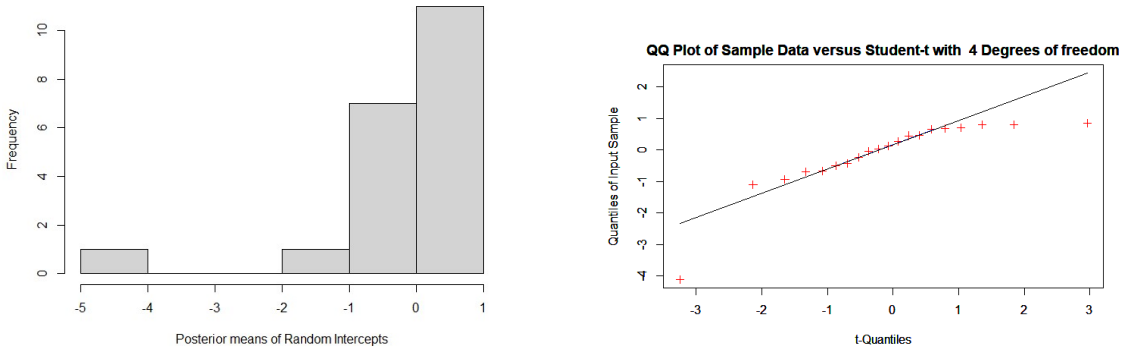


Figure 13: Histogram of the posterior means of the random intercepts (left) and Q-Q plot of the quantiles of the random intercept posterior means against the theoretical quantiles of the t-distribution (right).

3 Censored model

Evaluate the effect of the covariate grubs' size.

The basic setting of the Lognormal model proposed in the first part of this question was used again, but now considering the censored character of the observations. This means that instead of fixed death dates, intervals in which the death took place are now considered as the response. Sampling settings, such as sampled size, burn-in size, initial values and prior distributions were chosen similar as in the simplified setting.
The Gelman convergence test resulted in a estimated potential scale reduction factor almost equal to 1, for all parameters. The Geweke dynamic test resulted in most Z-scores within the two standard deviation range for both chains, as it can be seen in Figure 14 for the first chain. Figure 15 and Table 9 show that also in this model no clear effect of the size of the grubs on the expected survival time was observed.

Similarly to part 2, a Weibull model now considering the censored case was fitted to the data. The sampling setting for this case was the same as in part 2. Since

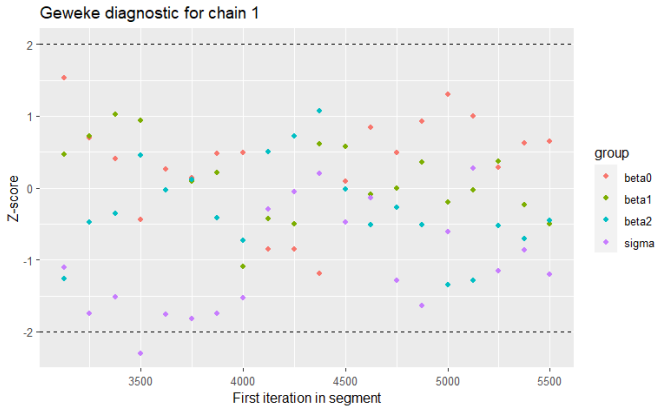


Figure 14: Geweke diagnostic for the first chain of the Lognormal model with right-censored observations.

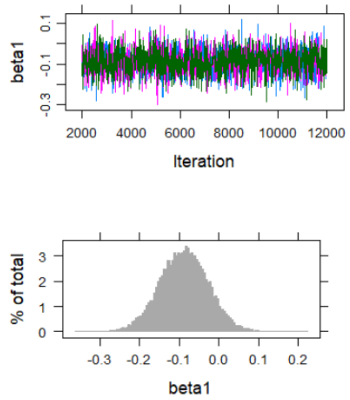


Figure 15: Traceplot and histogram of sampled values of β_1

	Mean	SD	2.5%	97.5%	MCSE
β_0	1.43	0.0848	1.26	1.59	0.00116
β_1	-0.0886	0.0612	-0.2089	0.0306	0.000496
β_2	-0.358	0.121	-0.595	-0.122	0.00164
σ	0.675	0.052	0.582	0.786	0.000453

Table 9: 95% credible interval for the model parameters

also the prior information with respect to all the observed data was the same as in part 2, the priors for the different parameters were also chosen identically. Again, the convergence was assured with formal tests. The Gelman test resulted in the estimated potential scale reduction factor equal to 1, for all parameters. The Geweke dynamic test resulted in most Z-scores within range for both chains, as it can be seen in Figure 16.

From Figure 17 and Table 10, it can be seen again that also in this model no clear effect of the normalized grubs' size was observed. As in part two, we have again

	Mean	SD	2.5%	97.5%	MCSE
β_0	1.77	0.091	1.59	1.95	0.00177
β_1	-0.0931	0.0671	-0.224	0.0392	0.000763
β_2	-0.427	0.127	-0.6777	-0.179	0.0024
scale	0.7091	0.054	0.613	0.825	0.000812

Table 10: 95% credible interval for the model parameters

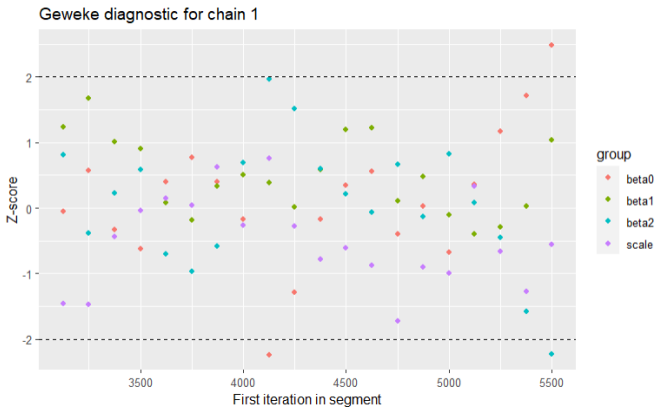


Figure 16: Geweke diagnostic for the first chain of the Weibull model with right-censored observations.

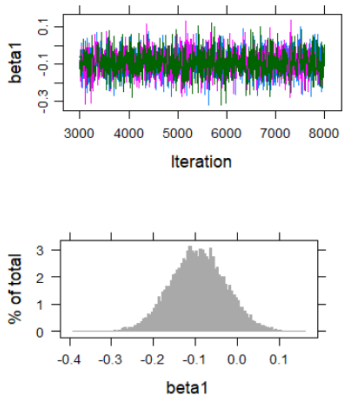


Figure 17: Traceplot and histogram of sampled values of β_1

compared the Lognormal model with the Weibull model in this part. For this we used the deviance information criterion as well as the pseudo Bayes factor again. In particular, the calculation of the DIC posed a challenge, since due to the censored nature of the data, there was no standard calculation method for the calculation of the effective number of parameters (p_d) of the model, see (Spiegelhalter et al. 2002) for an explanation. This problem was caused by the necessity to calculate the p_d value based on the deviance of the data, calculated on the basis of the mean estimate of the explanatory variables $D(\bar{\theta})$. However, since no exact survival times are available, there are no clear values for the survival time with which this mean estimate of the parameters could be compared, although these would be needed for this purpose. To solve this problem, we decided to use the mean sampled observed survival time to compare with the mean estimated parameters and thus to calculate p_d. We think that this is a reasonable way to solve the problem, since it accounts for the fact that survival times are not known and therefore we have to approximate them without making completely static assumptions, and the magnitudes of the estimated p_d values are at least kind of similar to those in part 2. There are probably other and perhaps better methods of approximation, but since the whole construct of DIC is somewhat heuristic, we think that the approximation we have made is sufficient. As in part 2, the DIC value of the lognormal censored model with 650.26 is again substantially lower than the Weibull censored model (DIC of 673.35). The value of the pseudo Bayes factor with a value of exactly 1.00 in the comparison of the two models, however, as in the last part of the work, provides no indication of which

model is to be preferred. If we compare the observed survival times again with the times estimated by the posterior predictive distribution, we find that the Weibull model estimates in about 10% of the cases at least as many survival times over 12 days as it was observed. For the lognormal model, it is about 11%. For values below 2 days, the lognormal model estimates about 81% and the Weibull model 97%. Therefore, the lognormal model again appears to perform slightly better overall. As

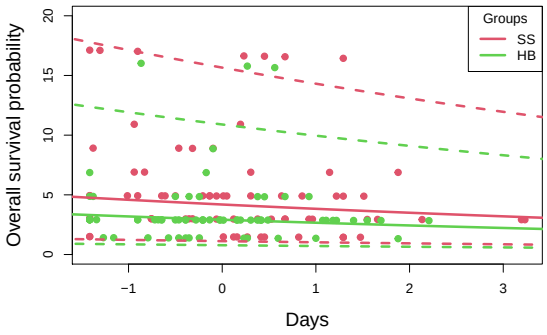


Figure 18: Estimated Survival Times

in Figure 6 in Figure 18[4] the estimated Survival times are pictured. One can notice that the mean estimates are almost identical, although in this case the estimated higher values for the data formerly constrained by 12 seem to be offset by the lower estimated values that are below 2.

Consider the clustering by including a random effect, i.e., the variation of location parameters across replicate plates. What does the random effect's variance component estimate suggest about the "clustering" of responses?

Lognormal and Weibull models with random intercepts are fitted also for the interval-censored problem. Results are similar with respect to the censored case without random intercepts. Once again, the grub size seems to be fairly irrelevant to the outcome, while grub type seems to be of greater importance. Additionally, like for the non-censored case, there seems to be an effect of the different plates on the size of the grubs. The convergence diagnostics were, like other models, indicative of convergence for all parameters. The sensitivity analysis also indicates invariance in the posteriors after varying priors (details can be found in the provided R-code).

[4]The points here represent the mean values of the Estimated Survival Time for each of the grubs.

	Lower95	Median	Upper95	Mean	SD	MCerr	MC%ofSD	SSeff	AC.10
σ^2	0.25	0.35	0.47	0.36	0.06	0.00	1.30	5730	-0.01
$\sigma^2_{b_0}$	0.04	0.18	0.43	0.21	0.12	0.00	2.40	1675	0.18
β_0	1.06	1.39	1.70	1.39	0.16	0.01	4.30	546	0.48
β_1	-0.17	-0.06	0.06	-0.06	0.06	0.00	1.30	5993	0.01
β_2	-0.76	-0.32	0.14	-0.31	0.22	0.01	4.30	544	0.48

Table 11: Summary of posterior parameters for Lognormal censored random-intercept model

Nevertheless, some of the settings led to the need for more burn-in iteration until the model had converged.
Also, the posterior predictive checks gave no evidence of a lack of fit as it was the case for the models in part 2
An AFT Weibull model including random intercepts was set up like the model with uncensored data, except now including censored data. The convergence diagnosis did not reveal any irregularities either. Results are shown in Table 12. Posterior predictive checks again show no irregularities in the distribution of random effects. The conclusions based on the posterior for γ are the same as for the noncensored model. For the censored Lognormal model with random intercept, the DIC value was 650.11, while the DIC value for the Lognormal model had a value of 627.27, which again favoured the Lognormal model over the Weibul model. Surprisingly, the pseudo-Bayes factor was 0.72 when comparing the two models, so it remains unclear which of the two models was a better fit to the data in this case. When simulating the posterior predictive distribution, it is again noticeable that high values in particular are reproduced rather poorly by both models, but especially by the Weibull model. Therefore, the conclusions of the model remain the same as in part 2 for the equivalent models.

	Lower95	Median	Upper95	Mean	SD	MCerr	MC%ofSD	SSeff	AC.100
γ	2.25	2.62	3.02	2.63	0.20	0.00	2.10	2324	0.07
β_0	3.14	4.01	5.04	4.03	0.48	0.01	2.70	1374	0.15
β_1	-0.42	-0.20	0.01	-0.20	0.11	0.00	0.90	13133	0.01
β_2	-1.70	-0.75	0.29	-0.75	0.50	0.01	2.20	1988	0.07
$\sigma^2_{b_0}$	0.15	0.92	2.27	1.06	0.63	0.01	1.30	5669	0.03

Table 12: Summary of posterior parameters for Weibull censored random-intercept model

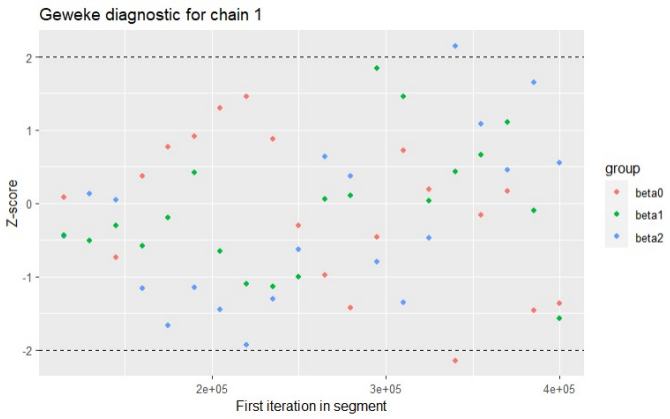


Figure 19: Geweke diagnostic for the first chain of the Weibull censored random intercept model.

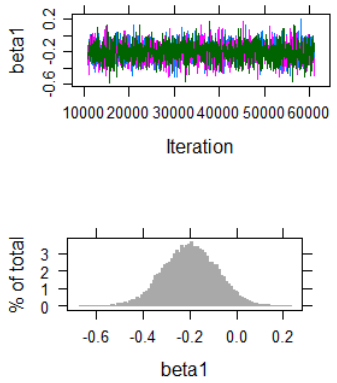


Figure 20: Traceplot and histogram of sampled values of β_1

Check whether there are outlying/influential observation(s). If so, give possible reasons for the outlying observation(s)

Conditional predictive ordinates (CPO) is used again to detect possible outliers of the log-normal model with random intercept: the estimate of the inverse-CPO is shown in Figure 21. We can see that, as in the non-censored case, the highest values refer to observations 91, 105 and 126, which are right-censored. The conclusions stay the same as in part 1. However these values are smaller than the non-censored case, showing that the model is less impacted by influential observations.

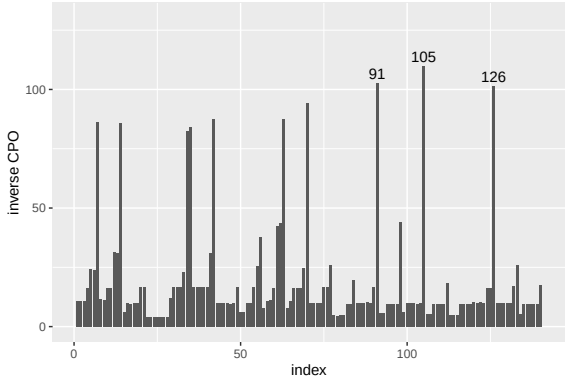


Figure 21: Index plot of inverse-CPO estimates

Check the distribution of the random effect and propose a robust distribution for the random effect.

Similar to the previous checks of the model based on the non-censored data, a histogram of the posterior means of the random intercepts and a Q-Q plot is shown in Figure 22. The plots are quite similar compared to the ones obtained when not considering the censored nature of the data. They show that normality is not followed closely and outliers are present. The PPCs shown in Table 13, show no evidence of a poor fit of the normal distribution for the random intercepts.

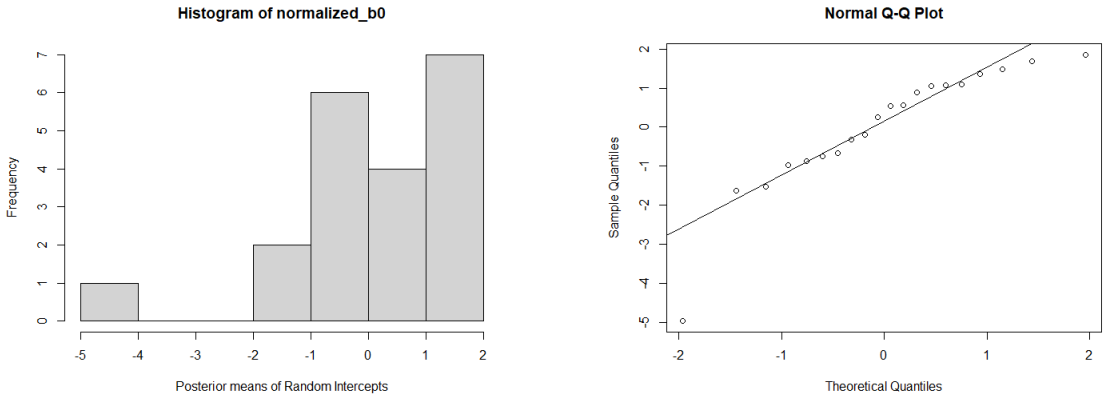


Figure 22: Considering censoring; Histogram of the posterior means of the random intercepts (left) and normal distribution Q-Q plot (right).

Statistic	p_D
Maximum	0.74
Minimum	0.75
Kolmogorov-Smirnov	0.33
Sinharay and Stern	0.82

Table 13: Posterior predictive p-values for various GOF statistics; censored data.

Once again, a t-distribution with low degrees of freedom (e.g. 4, as used previously), can be assumed for the distribution of the random intercepts, as a more robust distribution to outliers. The results are shown in Figure 23. The improvement in fit is once again not great, as significant deviation is observed in the right tail.

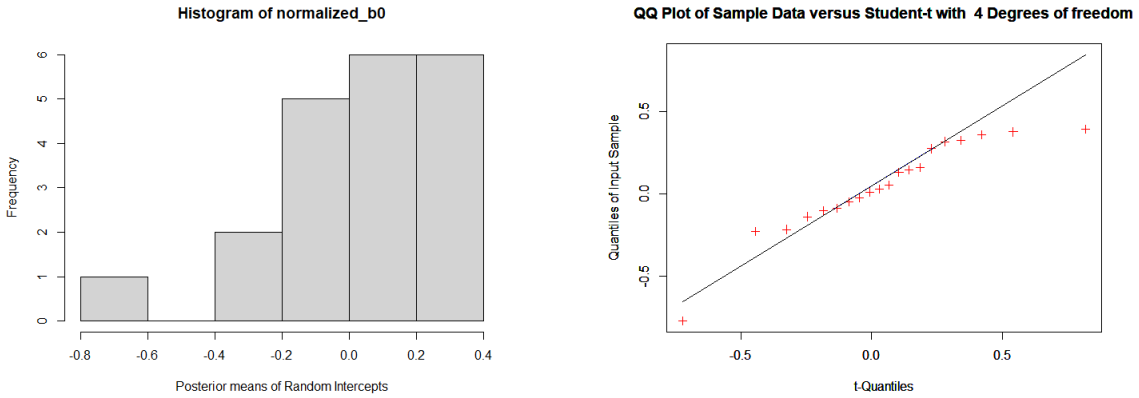


Figure 23: Considering censoring; Histogram of the posterior means of the random intercepts (left) and t-distribution Q-Q plot (right).

4 Conclusion

Various models and statistics were used to try to answer the initial research questions. Whether with or without random intercept and whether with or without censoring, the lognormal models on average described the observed survival data somewhat better than the Weibull models. However, the actual estimates are relatively similar for all models. Overall, it crystallizes in all cases that size does not play a clearly identifiable role in the estimation of survival times of individual EPNs, since none of the models assumed a value so far different from zero that the prediction interval no longer included zero. Moreover, if there is an effect, it is rather small. The estimated difference between the groups was on average about one day, which is clear enough that it can hardly be explained by chance. In the future it would be advisable to describe also other models than the present ones to examine the data, since in particular the many high values, although the models were not rejected by our analyses, nevertheless rather more improbably would occur. Overall, our work has fundamentally answered the questions formulated at the beginning but further analyses are necessary to come to a final conclusion.

References

Bayarri, M. J. and James O. Berger (2000). "P Values for Composite Null Models". In: *Journal of the American Statistical Association* 95.452. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.2000.10474309, pp. 1127–1142. ISSN: 0162-1459. URL: https://doi.org/10.1080/01621459.2000.10474309 (visited on 01/07/2022).

Diya, Luwis et al. (2012). "The relationship between in-hospital mortality, readmission into the intensive care nursing unit and/or operating theatre and nurse staffing levels". In: *Journal of advanced nursing* 68.5, pp. 1073–1081.

Lange, Kenneth L, Roderick JA Little, and Jeremy MG Taylor (1989). "Robust statistical modeling using the t distribution". In: *Journal of the American Statistical Association* 84.408, pp. 881–896.

Lesaffre, Emmanuel and Andrew B Lawson (2012). *Bayesian biostatistics*. John Wiley & Sons.

Spiegelhalter, David J. et al. (2002). "Bayesian measures of model complexity and fit". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64.4. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9868.00353, pp. 583–639. ISSN: 1467-9868. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00353 (visited on 01/09/2022).