Small sample considerations for anthelmintic resistance tests

---

Master Thesis in Biostatistics (STA495)

by

Tea Isler
16-718-850

supervised by

Prof. Dr. Reinhard Furrer

University of Zurich[UZH]

Zurich, April 2018

Abstract

The faecal egg count reduction test described in Coles *et al.* (1992) is the most common technique to asses anthelmintic resistance of parasites in horses. In the past years new statistical models have been proposed to find a solution to the limitiations of the FECRT. One of the latest is the zero-inflated Bayesian hierarchical model from Wang *et al.* (2017b) which is characterized by the ability to capture the high number of zero counts coming from unexposed livestocks. All these models though require a minimum number of 10–15 animals per stock which might often not be reached in practice.

In this thesis we focus on the problem of very small sample size using, for our analysis, the Bayesian hierarchical model from Wang *et al.* (2017b). We start with some considerations regarding point estimates and dilution factors showing that if the number of samples collected is less than 10, a dilution factor of maximum 20 should be used. We sequently discuss the influence and importance of an informative prior distribution for the reduction and implement two functions for the computation of the hyperparameters. A simple version of the hierarchical model is then proposed. With a simulation study we show how, for very small samples, the simple model performs better in terms of both bias and variance compared to the zero-inflated model. Final considerations are then made regarding the interpretation of the output where we suggest the use of posterior distributions to draw more reliable conclusions. Furthermore this thesis offers a schematical flowchart and a guide which we believe can help practitioners perform their analysis autonomously.

Aknowledgement

I would first like to thank my supervisor Prof. Dr. Reinhard Furrer for his advice and support. He was always open to my ideas and suggestions and helped me find new paths when the ones I was following seemed to have reached a dead end.

I would also like to thank Dr. Miriam Scheuerle and her team (`http://www.laborparadocs.de/`) for sharing the data used in this thesis, which gave me the possibility to have an insight on real life problems. In addition, I thank Craig Wang, for always being available to answer my questions, and Eva Furrer who was an important figure throughout my studies.

Finally I would like to thank my family and friends for their continuous support and encouragement. Without you I would have probably finished my studies but I would have gone insane. Thank you!

Tea Isler
Zurich, April 2018

Contents

Chapter 1

Introduction

1.1 Gastrointestinal nematodes and the McMaster technique

Gastrointestinal nematodes are one of the major types of livestock parasites infecting cattle, sheep and horses. These infections are a major cause in reduction of the profitability due to weight loss, reduction in milk and wool production. For example, a study in Australia has reported an estimated cost of 1 billion (Australian) dollars associated with parasitic diseases in sheep and cattle (McLeod, 1995).

The life cycle of gastrointestinal nematodes begins with adult female parasites laying thousands of eggs which then pass out in the feaces of infected animals. These eggs then develop to a first stage larva which hatches out of the egg in several hours. The process continues with the larvae feeding on bacteria and, by undergoing two molts, it finally reaches the third stage larva, which is infective to horses. These infective larvae migrate up wet blades of grass where they are most likely to be eaten by some animals, including horses (for a more detailed description see Boehringer Ingelheim (2017)).

Currently on the market we can distinguish three classes of anthelmintic drugs for horses: benzimidazoles (fenbendazole, oxibendazole), tetrahydropyrimidines (pyrantel) and macrocyclic lactones (ivermectin, moxidectin). The latter ones are the most effective compounds against cyathostomins (group of small strongyles), while resistance to tetrahydropyrimidines and specially benzimidazoles is quite common and increasingly reported (Sanna *et al.*, 2016).

The most used and efficient method to diagnose parasitic gastroenteritis and determine the efficacy of anthelmintic treatments is faecal examination. There are several technique to count the number of eggs in feaces but the most common one is the modified McMaster method by Coles *et al.* (1992), which follows the following steps:

1. a small amount of feaces is collected and diluted with a chosen flotation fluid;
2. the solution is stirred to ensure homogeneous distribution of the eggs and filtered;
3. a Pasteur pipette is used to withdraw 2 sub-samples (one for each chamber of the McMaster slide);
4. the chamber is let stand to allow the eggs to float to the surface;

5. with the use of a microscope, the number of eggs inside the grid area of the chamber are counted.

To obtain the faecal egg count (FEC), the number of eggs observed in the McMaster chamber is multiplied by analytical sensitivity f, which is calculated as the dilution factor divided by the volume of the McMaster chambers and the grams of feaces analysed (we will talk about it in more detail in Section 2.4).

1.2 The faecal egg count reduction test

The first evidence-based method to detect a reduction in faecal egg counts (FEC) is the so called "faecal egg count reduction test" (FECRT). It was suggested in the The World Association for the Advancement of Veterinary Parasitology (WAAVP) and because of its simplicity and easy interpretation it is the most common method used worldwide. The FECRT can be performed on both paired and unpaired data, but throughout the thesis we will only consider the paired case where the same horses are tested twice, once before starting the treatment (day 0) and once after (usually 7, 14 or 21 days later, depending on the drug). The egg count reduction percentage (FECR%) after 14 days of treatment is then obtained from the following expression:

$$\left(\frac{\mathrm{FEC}_0 - \mathrm{FEC}_{14}}{\mathrm{FEC}_0}\right) \cdot 100$$

where FEC_0 and FEC_{14} are the arithmetic means of FEC of all animals at days 0 and 14 respectively.

Furthermore, to give a more accurate indication of the range of the data (Lester *et al.*, 2013), 95% confidence limits were included by using bootstrapping from the observed FECR% and then taking the upper and lower 2.5-quantiles of N simulations as the 95% confidence limits (Torgerson *et al.* (2005), Vidyashankar *et al.* (2007)).

The classification of FECRT

According to Coles *et al.* (1992), there are three possible outputs for the FECRT when studying sheep and goats:

- resistance is present: the FECR% in egg counts is less than 95% *and* the 95% lower confidence level (LCL) is less than 90%
- resistance is suspected: if only one of the two criteria mentioned above is met
- inconclusive: neither of the above criteria are met

Unfortunately no such defined guidelines regarding the cut-off values are present for horses. Coles *et al.* (1992) only reported that a reduction of 90% or less is indicative for benzimidazole (BZD) resistance. Further

cut-off limits were suggested by Kaplan and Nielsen (2010) as 90% for pyrantel (PYR) and fenbendazole (FBZ) and 95% for ivermectin (IVM) and moxidectin (MOX). In their paper however Kaplan and Nielsen (2010) use a binary classification for the output of the reduction test (resistant or suspected) without considering confidence intervals. In recent studies such as Relf *et al.* (2014), Lester *et al.* (2013) and Fischer *et al.* (2015) the threshold for the 95% lower confidence level was set to 80% for FBZ and PYR and 90% for IVM and MOX. The upmentioned trichotomization have several limitations, not only it is easier to misclassify the outcome when it is closer to the thresholds, but it also does not provide a complete picture of the output, which may bring to misleading conclusions. In Chapter 2 we will see how the posterior probabilities from Bayesian models can provide a more accurate interpretation of the outcome, specially when working with small samples.

1.3 From the FECRT to the ZIP model

In the past years, several limitations of the FECRT have been identified. The FECRT does not account for the variability of FEC data and aggregation or overdispersion of the egg counts distribution. A first attempted solution for the latter problem was proposed by Torgerson *et al.* (2005) by assuming a negative-binomial distribution for the counts and by using a bootstrap method to compute confidence intervals of FEC reduction. A second idea came from Vidyashankar *et al.* (2007) who proposed a non-parametric bootstrapping method, which re-samples and summarises the observed data without making assumptions regarding the underlying distribution. The limitation of this method though, is the assumption that the observed data is fully representative of the population, which is most likely not the case when working with small sample sizes. The third method, proposed by Denwood *et al.* (2010), makes use of the Markov Chain Monte Carlo and assumes two different gamma-Poisson (negative binomial) distributions for pre- and post-treatment data. This model accounts for FEC data variability but still ignores the high frequency of zero observations, which are due to unexposed animals. Wang *et al.* (2017b) presented a zero-inflated hierarchical model based on the work from Paul *et al.* (2014). The zero-inflated Poisson model (ZIP) described in the paper not only performs better than all the upmentioned techniques, but also presents ad-

ditional advantages such as density distributions of the model parameters and flexibility in model formulation (such as adjustments to animal-specific reductions).

1.4 Thesis motivation and contribution

All the methods mentioned in the provious section are based on a sample size of 10 or more animals raising the question on how to deal with smaller livestocks.

In this thesis we make some considerations on such small samples and try to define a set of guidelines which can be followed by practitioners. More specifically, we will work with some of the options available in the `eggCounts` package (Wang *et al.*, 2017a) such as analytical sensitivity, prior distributions, etc. to reduce the inevitable uncertainty that arises when analyzing very small datasets. Further improvements are tried to achieve by simplifying the Bayesian hierarchical model and by restricting the data to a paired design and including only observations with initial FEC higher than 200 eggs per gram (measurement unit used in anthelmintic resistance tests).

For a better overview, the thesis contributions are split into three parts.

Theoretical considerations:

- analysed the effect of the analytical sensitivity f on the output;
- highlighted considerations regarding the McMaster, FLOTAC and mini-FLOTAC techniques;
- discussed the influence of an informative prior for the reduction distribution δ;
- pointed out the differences of working with and without truncated data to raise awarness;
- proposed new simple model for small samples;
- recommended the use of the posterior distribution to interpret the output.

Shiny web interface:

- implemented options for the use of an informative prior for the reduction parameter;
- added options regarding samples size to allow the use of the simple model.

`eggCounts` **package:**

- added function `getPrior_delta` to compute the hyperparameters of the prior distribution for reduction;
- added a new `fecr_stanSimple` function for the analysis of small samples.

Chapter 2

The eggCount package for small samples

2.1 Considerations on small sample size

It is well known that in statistics, and specially in frequentist inference, most of the key concepts as the power, confidence interval and effect size are all affected by the number of samples included in the study. A very small sample size, as in our case, produces high variance and low reproducibility. A better approach to such problem is Bayesian inference, where information from prior experiments can be incorporated in the analysis of current datasets. Of course we recognize that Bayesian inference is not going to solve all our problems and that it is inevitable that a small dataset with four to ten subjects is still going to produce high uncertainty.

When the `eggCounts` package was first developed (version 0.2) by Torgerson *et al.* (2014), a sample size of ten animals was used in the simulation study. The results showed insufficient evidence to conclude that resistance was present and this might have been due to a low pre-treatment egg counts or a too small sample size. In this thesis we will use the latest version of the `eggCounts` (Wang *et al.* (2017a)) which still recommands a minimum sample size of 10–15 animals. We will simulate and analyse datasets containing a very low number of subjects (from four to ten animals) and make considerations on the outputs. Because of such a small sample, throughout the thesis we will only work with the non-zero-inflated paired model. The goal of this chapter is to try to reduce the large uncertainty of the outcome by using the numerous variables available in the package.

2.2 The Bayesian hierarchical model

The version of the `eggCounts` package described in Wang *et al.* (2017b) is based on the Bayesian hierarchical model suggested by Paul *et al.* (2014). Intuitively, the word "hierarchical" indicates that the Bayesian model is in a hierarchical form, meaning that the parameters of the prior distributions are not fixed but dependent, in turn, on additional parameters. These last parameters take the name of hyperparameters. The hierarchical model for the paired design proposed by Paul *et al.* (2014), given the true number of eggs per gram (epg) of faeces Y_i^b, considers a binomial distribution for the observed number of epg Y_i^{*b} with probability p and size Y_i^b. The p is the proportion of diluted faecal suspension that is put on a McMaster slide and it is obtained from the expression $1/f$, where f is the analytical sensitivity used in the McMaster method. Given then the mean number of epg of feaces μ_i^b, the Y_i^b are assumed to follow a Poisson distribution with mean μ_i^b. In turn, μ_i^b is left variable in order to account for heterogeneity between animals, hence a Gamma distribution with shape parameter ϕ and rate parameter ϕ/μ is assumed. What we just described is the pre-treatment model (the letter b stands for "before"). In formula this leads to:

$$\begin{aligned} Y_i^{*b}|Y_i^b &\sim Bin(Y_i^b, p) \\ Y_i^b|\mu_i^b &\sim Po(\mu_i^b) \\ \mu_i^b|\phi, \mu &\sim Gamma(\phi, \phi/\mu). \end{aligned}$$

After the treatment period (7, 14 or 21 days) the number of eggs are counted again and the epg rate is reduced by a factor $1 - \delta$. The model after the treatment is then

$$\begin{aligned} Y_i^{*a}|Y_i^a &\sim Bin(Y_i^a, p) \\ Y_i^a|\mu_i^b, \delta &\sim Po(\delta\mu_i^b). \end{aligned}$$

Notice that the mean μ_i^b is the same in both models. This is to indicate that the pre and post treatment epg Y_i^b and Y_i^a belong to the same animal.

The μ, ϕ and δ are hyperparameters and appropriate distributions must be assigned. The parameter μ is the global mean epg for infected horses and Paul *et al.* (2014) proposed a $Gamma(1, 0.001)$ distribution where 90% of the prior probability mass lies essentially between 50 and 3000. For the dispersion parameter ϕ a prior distribution $Gamma(1, 0.7)$ is assumed.

Here 90% of the prior probability mass lies between 0.1 and 4.3. Finally, the reduction parameter δ follows a $Beta(1,1)$ which is equivalent to the $Unif(0,1)$ distribution.

2.3 Simulation and analysis of small datasets with the eggCounts package

In this section we will see how the `eggCounts` package performs when analysing small datasets. We will simulate a paired design dataset with 4 subjects by using the function `simData2s()`. This function takes into account several parameters including the true number of epg before treatment, the proportion of epg after treatment, the dispersion parameter, the pre- and post-treatment prevalences and the analytical sensitivity. The FECs reduction is then modeled by the function `fecr_stan()` which uses the Stan modelling language.

```
counts <- simData2s(n = 4,          # number of samples
                    preMean = 150, # samples mean
                    delta = 0.1,    # 1-delta=reduction
                                    # -> reduction = 90%
                    kappa = 1,      # dispersion parameter
                    phiPre=1,       # prevalence
                    f=50)           # analytical sensitivity
resultsP <- fecr_stan(counts[,"obsPre"], counts[, "obsPost"],
                      zeroInflation = FALSE)
```

The `fecr` object in the output is the reduction we are interested in. We know that the true reduction is 90% and want to check if the results obtained from the model correspond.

Since the `fecr_stan()` performs Bayesian statistical inference, the output will not be a single reduction value but rather a table including the mean, median and mode together with the most important quantiles and the high and low 95% highest posterior density (HPD) from the 4000 iterations (the default number of iterations in `fecr_stan()`). Because each simulation produces a different dataset (hence a different output from the model), it is important to have an overview of the behaviour of the point estimates and see how well the model performs. We here simulate 1000 datasets and analyse the outcome.

```
round(resultsP$posterior.summary, 2)

##                        mean     sd  2.5%    25%    50%    75%
## fecr                   0.78   0.14  0.41   0.71   0.81   0.89
## meanEPG.untreated    270.64 191.74 85.72 158.31 218.88 315.36
## meanEPG.treated       56.29  63.74  7.78  24.67  40.47  64.43
##                       97.5% HPDLow95   mode HPDHigh95
## fecr                   0.96     0.50   0.88      0.99
## meanEPG.untreated    818.05    54.05 165.81    652.15
## meanEPG.treated      192.54     1.53  28.91    148.46
```

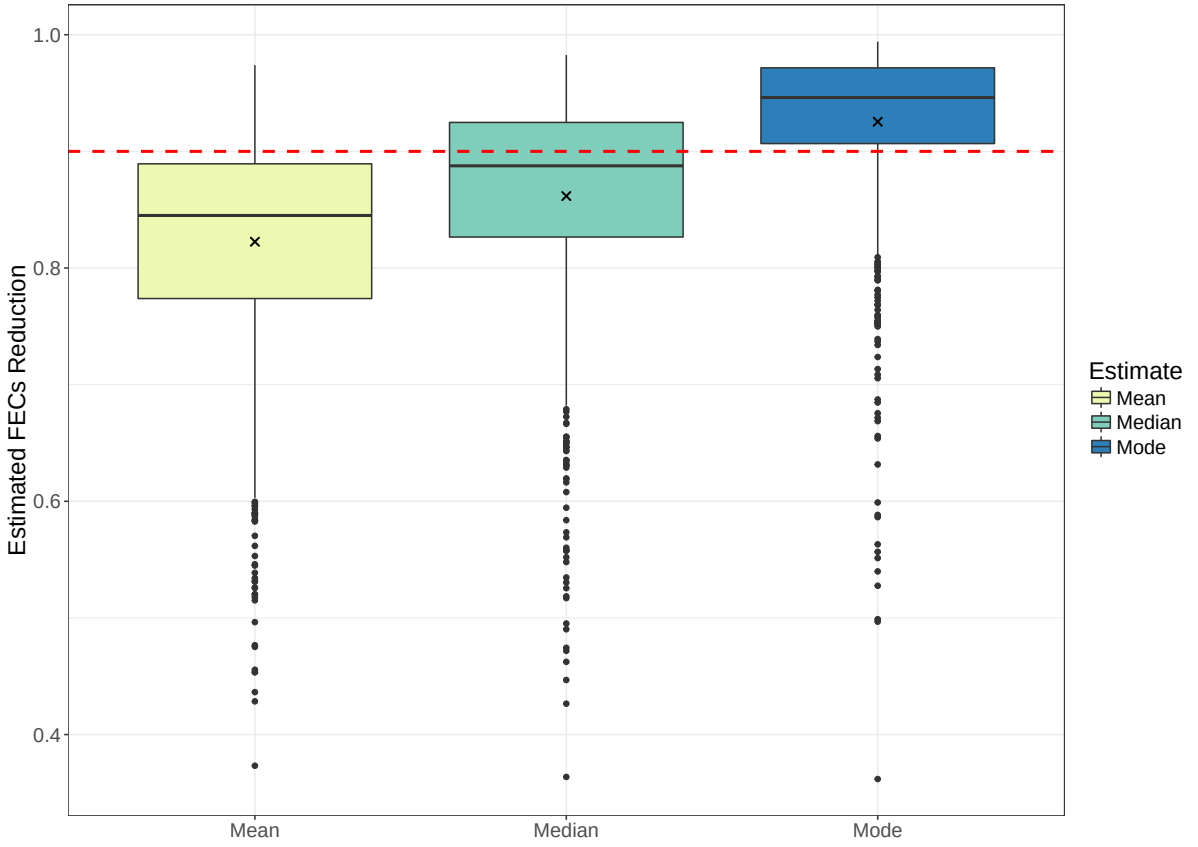

Figure 2.1: Distribution of 1000 simulations of epg reduction after treatment. The red dashed line represents the true reduction of 90% and the black cross indicates the mean.

Figure 2.1 gives a graphical representation of the three point estimates for the reduction mean obtained after repeating the analysis 1000 times. Each black point rfnot only represents an iteration and the red dashed dashed line indicates the true reduction which was specified in the simulation process.

We can observe that the posterior mean point estimate underestimates the reduction, hence we will not be considering it anymore. As analysed in Wang *et al.* (2017b), also here we observe a lower bias for the mode estimate compared to the median but since the latter might sometimes be prefered (i.e., when the distribution for the data is skewed), in the next section we will still report both quantities.

2.4 What can be done?

In this section we will make some considerations regarding the parameters involved in the hierarchical model used in the `eggCounts` package. By simulating datasets we will observe that appropriate modifications of some parameters will lead to a lower variance of the FECs reduction.

2.4.1 The analytical sensitivity f

The first and most natural thing to do is to improve the sensitivity of the test. The analytical sensitivity, expressed by f, is calculated by dividing the dilution factor by the volume of the McMaster chambers and the grams of feaces analysed. Hence a test with a two-chambered slide and a dilution factor of 75ml that analyses 5 grams of feaces would give a sensitivity of $f = 75/(5 \cdot 2 \cdot 0.15) = 50$. This value is then used to compute the number of eggs per gram of faeces (epg):

$$\text{epg} = \text{observed number of eggs} \cdot f.$$

In the expression for the analytical sensitivity the only parameters we can vary are the dilution factor and the grams of feaces, since the volume of the McMaster chamber is fixed to 0.15ml. A higher sensitivity, typically $f = 15$, provides a better approximation of the true number of eggs, although, because of the random distribution of eggs within a faecal sample, there is inevitable variability in evaluating the true number of eggs.

Figure 2.2 and Figure 2.3 give an overview of the estimated FECs reduction with different values for the analytical sensitivity f. For a baseline mean count of 150 epg, it can be seen that, when choosing the median as point estimate, a higher sensitivity (smaller f) performs better in terms of bias and slighlty better in terms of variance. On the other hand, such improvement is not as evident when working with the mode point estimate.

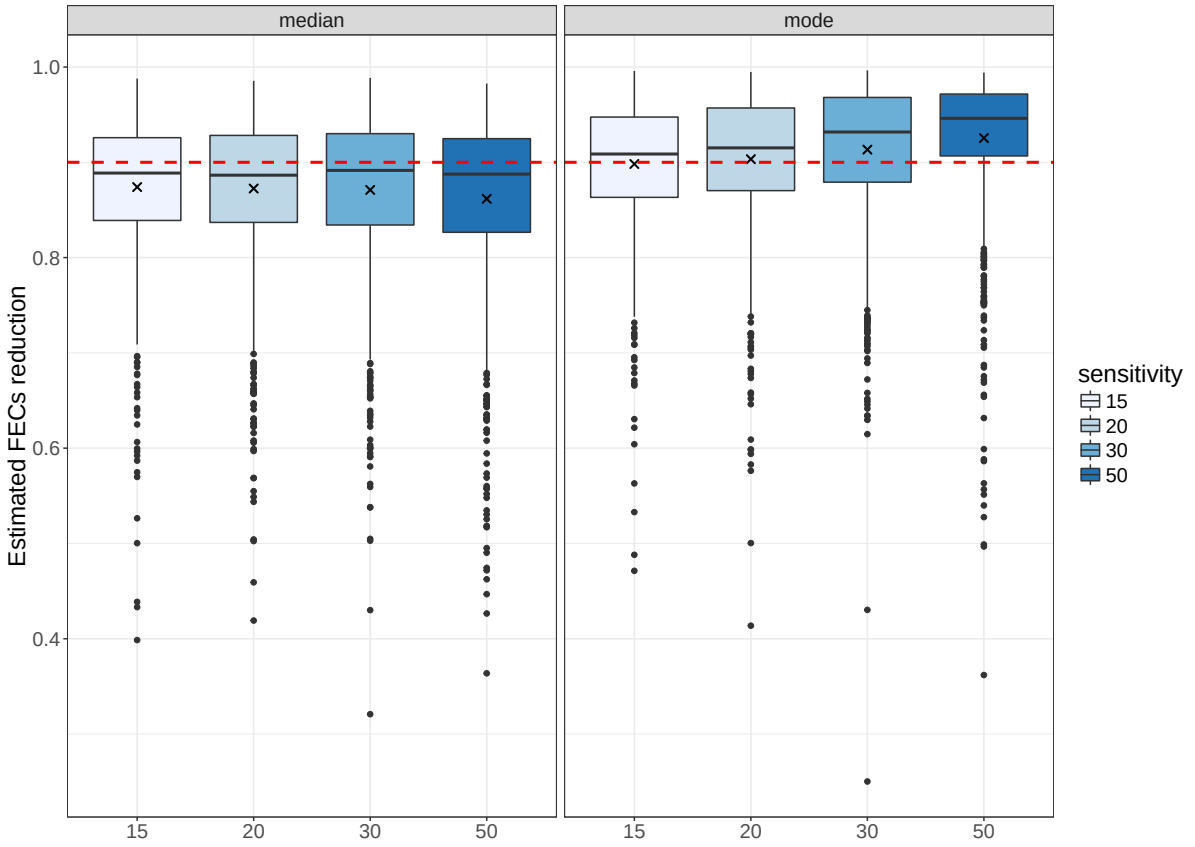

Figure 2.2: Two boxplots of the estimated FECs reduction with an initial true number of epg of 150 and an analytical sensitivity f of 15, 20, 30, 50. The one on the left uses the median as point estimate, while the one on the right uses the mode. Each black point represents one of the 1000 iterations while the red dashed line is the true reduction 90% specified when simulating the datasets. The black crosses indicate the mean.

Except for two outliers for f=50, we cannot see considerable improvements with higher sensitivities. A slightly different conclusion can be seen when the initial true number of epg is increased to 500. For both estimates, the estimated reduction is almost never less than 50%, meaning less variance compared to the case with initial epg of 150. Here the difference between the two point estimates is more evident. The mode performs better for all four cases in terms of bias, although such improvement is not so pronounced in terms of variance.

To summarize, when dealing with a small sample (here 4 subjects), we suggest to work with an analytical sensitivity of 20 or less. For these sensitivities, the mode point estimate performs better than the median in both cases of low and high number of epg before treatment. In the case of a higher number of epg before treatment (here 500) the mode point estimate provides lower bias and variability compared to the case were μ is small.

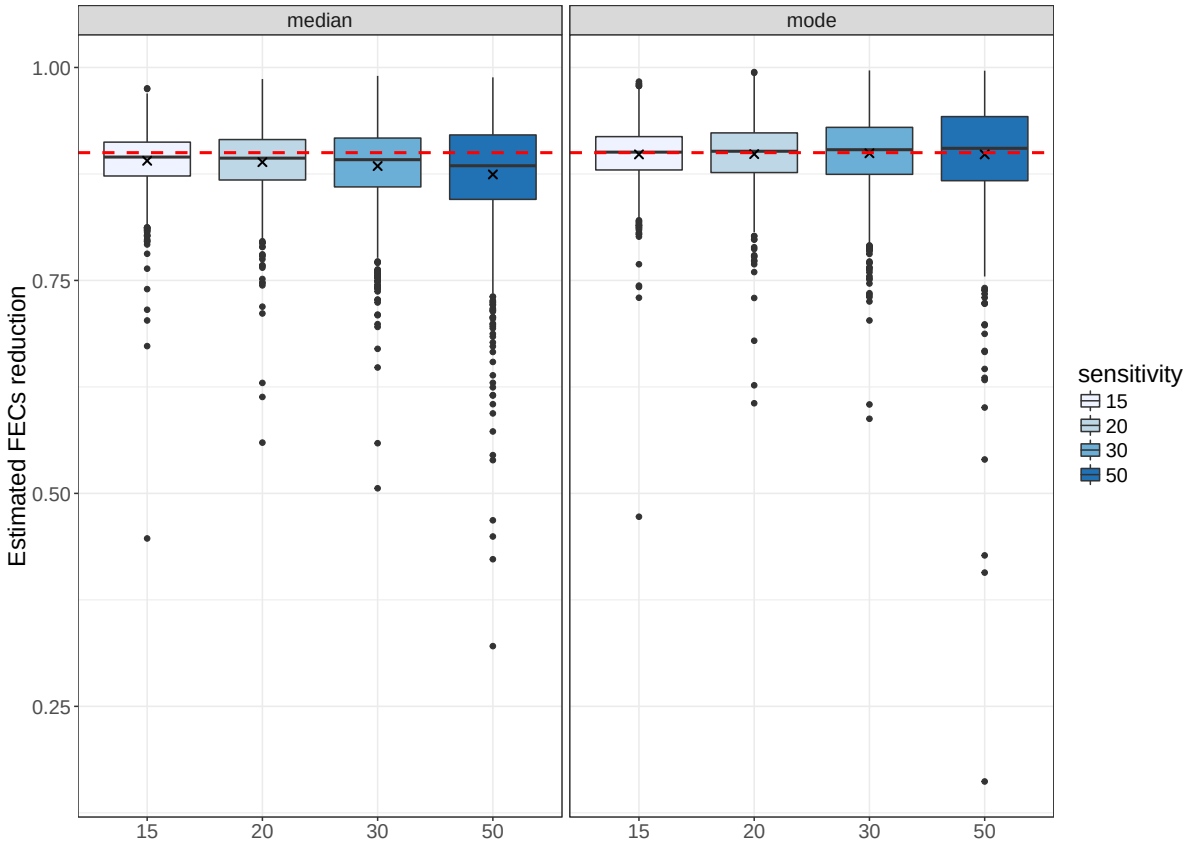

Figure 2.3: Two boxplots of the estimated FECs reduction with an initial true number of epg of 500 and an analytical sensitivity f of 15, 20, 30, 50. The one on the left uses the median as point estimate, while the one on the right uses the mode. Each black point represents one of the 1000 iterations while the red dashed line is the true reduction 90% specified when simulating the datasets. The black cross indicates the mean.

2.4.2 FLOTAC and Mini-FLOTAC

Until now, we have analyzed the model with different values for the analytical sensitivity when using the McMaster technique. This technique is the most widely known and used because of its simplicity and affordability, but it lacks in sensitivity. When using the McMaster technique the highest analytical sensitivity we can reach is approximately 10–15 epg. In Cringoli (2006) the FLOTAC technique was introduced to overcome the low sensitivity issue. FLOTAC is based on the centrifugal flotation of the sample and translation of the floating suspension on the top layer. Unfortunately, despite the accuracy and high analytical sensitivity of 1 epg, this new apparatus introduces complexity due to centrifugation, which requires specific laboratory equipment. The new Mini-FLOTAC technique described in Cringoli *et al.* (2013) is a simplified technique which can be

more easily carried out in laboratories with limited facilities. The analytical sensitivity of the Mini-FLOTAC is 5 epg.

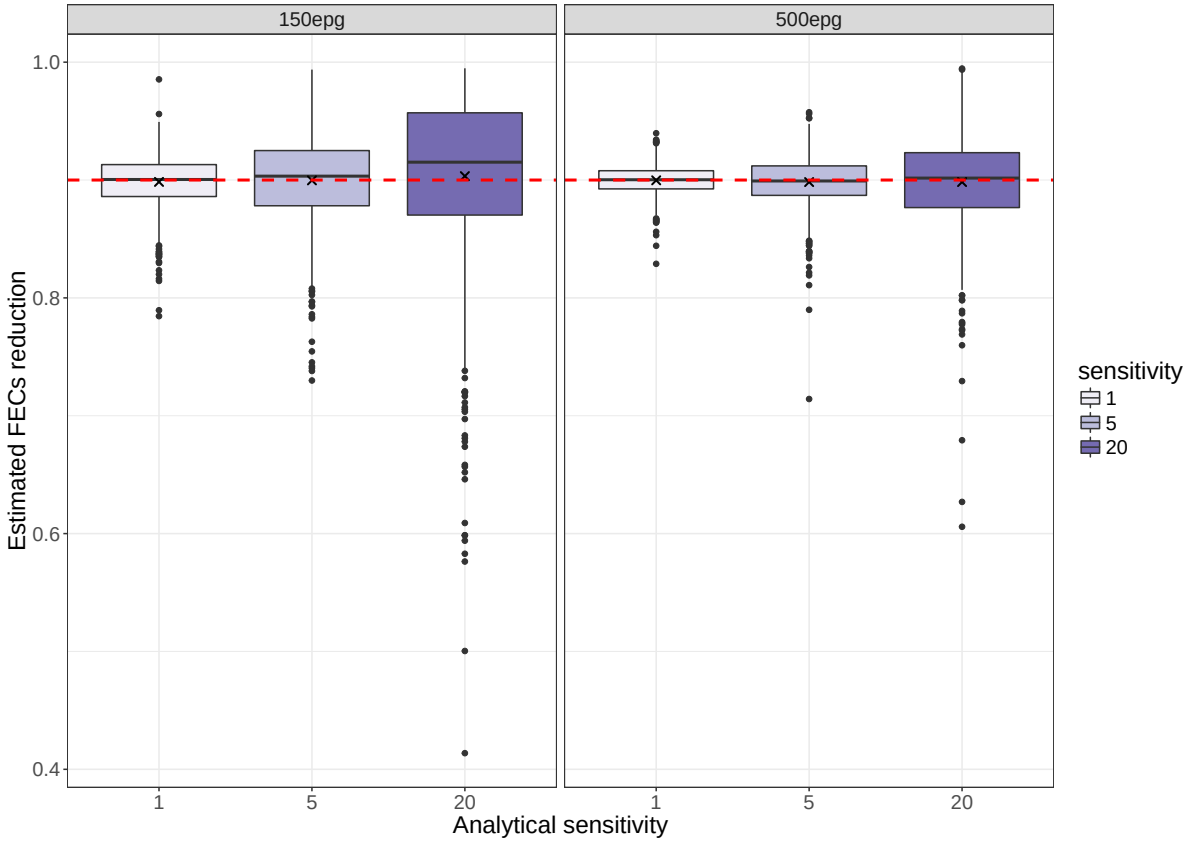

Figure 2.4: Comparison in terms of sensitivity between FLOTAC, mini-FLOTAC and McMaster (with f=20) techniques. In the left plot we used an initial true mean of 150epg to simulate the data, while for the right plot the number of eggs was raised to 500.

We now simulate 1000 datasets with a sensitivity of 20 (McMaster), 5 (Mini-FLOTAC) and 1 (FLOTAC). The comparison is meant to briefly show the significant difference between the techniques and arise awarness on the importance of the technique choice (for a more detailed analysis see Noel *et al.* (2017)). Figure 2.4 shows the 1000 mode point estimates from the posterior distribution of the reduction parameter. It is evident that as the analytical sensitivity increases (smaller f), the uncertainty reduces drastically, providing more accurate estimates of the true reduction. The effect is particularly strong when the initial mean number of epg is low (left panel of Figure 2.4)

We realize that the FLOTAC technique might be complex and expensive, but we strongly recommand practitioners to consider using the simpler Mini-FLOTAC technique which nevertheless provides lower uncertainty

compared to the McMaster technique in both cases of low and high number of epg before treatment.

2.4.3 The reduction parameter

At the beginning of this chapter we briefly pointed out the advantage of performing Bayesian analysis when working with small sample size. An informative prior distribution might solve the issues of low power and biased parameter values. The question that arises is: how to choose an informative prior? In van de Schoot *et al.* (2015) they listed three steps to specify a prior distribution:

- research and background knowledge (such as meta-analysis or previous comparable studies),
- choice of the type of distribution (normal, Poisson, etc.) and
- choice of values for the hyperparameters (shape and rate, mean and variance, etc.).

In this thesis we are interested in the reduction of epg, for which a non-informative prior distribution Beta(1,1) is specified in the `eggCounts` package. Because the Beta distribution is defined on the interval [0,1] and allows for different probabilities for each value in the interval (contrarily to the uniform distribution), there is no need for background knowledge to agree that the Beta is the right type of distribution for our scenario. The previous question would now become: how to choose appropriate hyperparameters? The answer is heavily dependent on background knowledge and must be assesed by performing a so called "sensitivity analysis".

Sensitivity analysis

As discussed by van de Schoot *et al.* (2015), the choice of specific hyperparameters for the prior distribution might affect the final output. When working with small datasets where more weight is given to the prior distribution, it is particularly important to analyse how sensitive the conclusions are with respect to the specified values of the hyperparameters. In this section we will perform a sensitivity analysis on the prior distribution of the reduction parameter δ. The choice of shape and rate parameters for a beta distribution is not as intuitive as it could be for the mean and variance of

a normal distribution. For this reason we decided to allow for two different methods for the computation of the Beta parameters. The first one is the reparametrization from Kruschke (2014) where shape a and rate b are replaced by mode ω and concentration k using the equations below:

$$a = \omega \cdot (k-2) + 1 \qquad \text{and} \qquad b = (1-\omega)(k-2) + 1,$$

where k and ω are real numbers with domains $[2, \infty)$ and $[0, 1]$ respectively.

The function in R to compute the parameters using this reparametrization is reported below.

```
shapes_Beta_OK <- function(omega, k){
  alpha <- omega*(k-2)+1
  beta <- (1-omega)*(k-2)+1
  param <- cbind(alpha, beta)
  colnames(param) <- c("alpha", "beta")
  return(param)
}
```

While the concept of mode is of general knowledge and widely used, the concentration value can be very abstract and subjective. It can be seen as a type of variance specifying how certain we are that our mode is the true population mode and where a low k indicates uncertainty. A mode of 0.5 and concentration of 2 (the lowest value k can take) correspond to the non informative prior $Beta(1, 1)$.

The second method to compute the parameters of the Beta distribution is based on the values of a suggested confidence interval for the reduction parameter, i.e. a practitioner believes the true reduction lies between 0.5 and 0.9 with a probability of 95%. In this case a and b can be computed in R using the function below, where x and y are respectively the lower and upper limits of the 95% confidence interval.

```
library(rootSolve)
shapes_Beta_CI <- function(x, y, p=0.95){
  # x = 2.5% quantile
  # y = 97.5% quantile
  # p = probability, area "most likely"
  f <- function(k){c(F1 = qbeta((1-p)/2, k[1], k[2])-x,
```

```
                    F2 = qbeta((1-p)/2, k[1], k[2],
                               lower.tail = FALSE)-y)}
  ss <- multiroot(f = f, start = c(1, 1))
  param <- rbind(ss$root)
  colnames(param) <- c("alpha", "beta")
  return(param)
}
```

To conduct the sensitivity analysis we work with the reparametrization mode/concentration. We choose 4 different modes ($\omega = 0.05, 0.2, 0.35, 0.5$) and 4 different concentration ($k = 5, 10, 15, 20$) for the prior distribution of the reduction parameter δ. Figure 2.5 gives a graphical overview of such distribution for each of the 16 combinations.

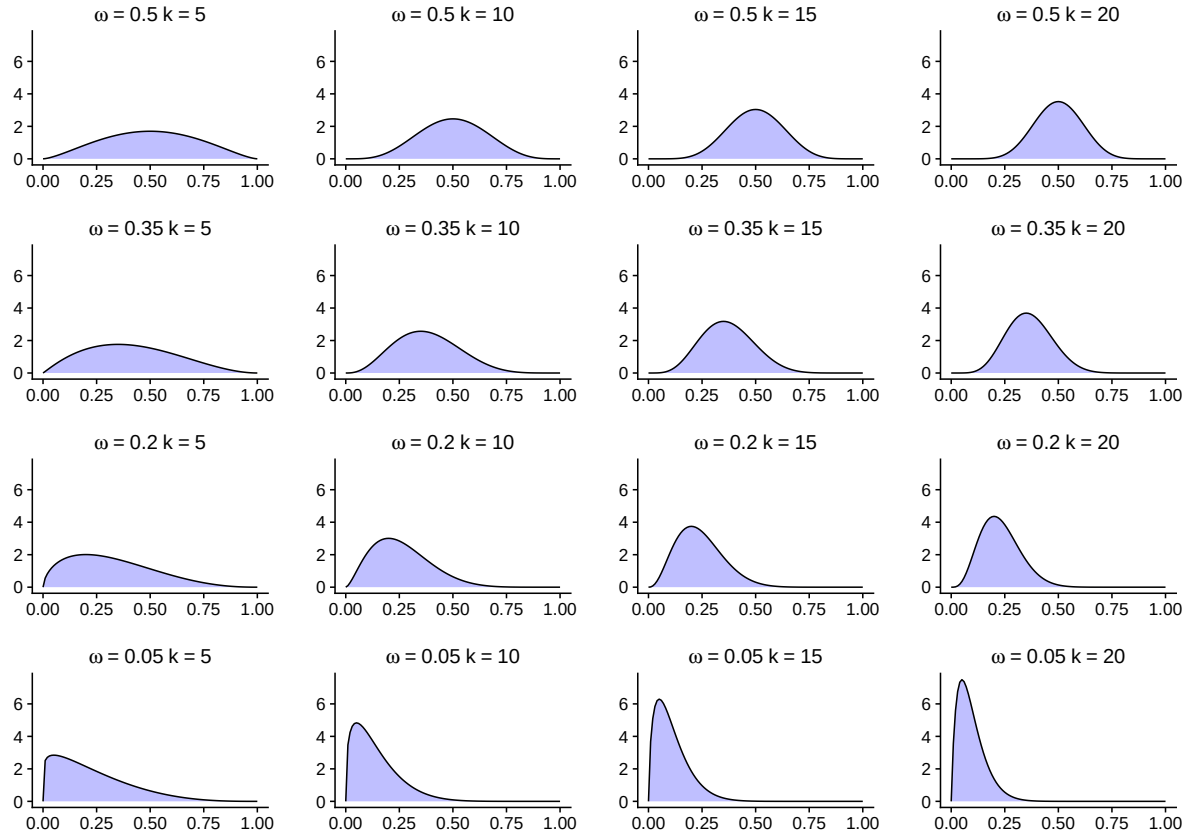

Figure 2.5: Graphical representation of the prior Beta distribution for each of the 16 combinations of mode and concentration. From top to bottom we observe the distribution shifting towards its mode while from left to right, as the concentration increases, the distribution tends to concentrate more around its mode.

The next step is to run the model for each combination of mode and

concentration on the same simulated datasets. Here, as before, we simulate and analyse 1000 datasets to have a more accurate picture of the behaviour of the model. The datasets are simulated using the `simData2s` function and consist of four observations (animals). For the parameters we choose an initial true number of epg of 250, a true reduction of 0.7 (meaning a δ of 0.3), a dispersion value of 1 and a dilution factor of 20. Furthermore we assume all animals are infected before treatment hence the pre-treatment prevalence is 1.

```
set.seed( 12)
counts <- simData2s(n = 4,  # number of samples
                    preMean = 250, # samples mean
                    delta = 0.3,  # 1-delta=reduction = 70%
                    kappa = 1, # dispersion parameter
                    phiPre=1,  # prevalence
                    f=20)    # dilution factor
counts

##      obsPre masterPre truePre obsPost masterPost truePost
## [1,]    240        12     225     100          5       84
## [2,]    100         5      75      20          1       28
## [3,]    140         7      99      20          1       30
## [4,]      0         0       4       0          0        0
```

The results of this sensitivity analysis are displayed in Figure 2.6. The figure shows the 1000 posterior reductions δ for each of the 16 combinations of mode and concentration. It can be seen that, as the concentration increases, the variance decreases and there is a shift of the distribution towards the true mode ω. Overall we can observe that the model is quite sensitive with respect to the prior distribution for reduction, which makes sense given the small sample analysed. We therefore recommand practitioners to carefully choose appropriate hyperparameters based on their knowledge and prior information. Alternatively, if no background is available, the non informative prior $Beta(1, 1)$ should be used.

2.4.4 Truncated data

Another aspect worth talking about is the tendency of truncating data. In real life, many practitioners consider an initial number of eggs of less than

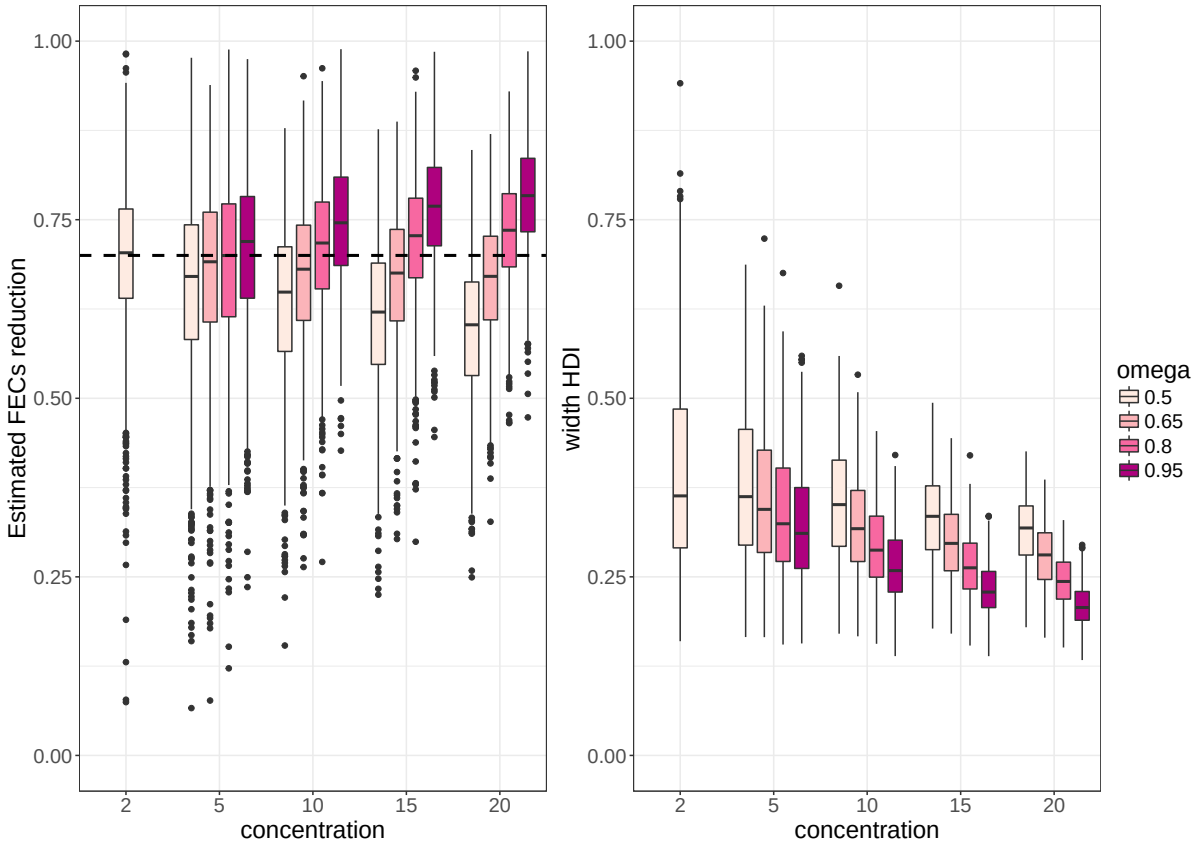

Figure 2.6: Boxplots of the 1000 posterior reductions (left) and widths of HPDIs (right) for each of the 16 combinations of mode ω and concentration k. The single boxplot with a concentration of 2 represents the distribution of the posterior reduction and HPDI width using a non informative prior $Beta(1, 1)$.

200epg to be low and would exclude the animal from the study. In this section we look at the difference when analysis only data with an initial epg of 200 or more and when instead all animals are treated, regardless of the number of epg before-treatment. The goal of this section is not to conclude which case is the best, but rather to raise awarness about the interpretation of the outcome of truncated data.

For each case (truncation/no truncation) we simulate 1000 datasets with a before-treatment mean of 250epg, a reduction of 70% and an analytical sensitivity of 20. Furthermore we consider no overdispersion ($\kappa = 1$) and a prevalence of 100%. For the non-truncated case we simply simulate datasets with 4 observations as we did up to now. To simulate the truncated data we instead generate 1000 datasets with 10 observations and then work with only the first 4 that have an epg larger than 200. The default model from the `eggCounts` package is then fitted to both cases of simulated datasets. The 1000 mode point estimates from the posterior distributions

of the reduction parameter are plotted on the left side of Figure 2.7. It can be seen that the truncated data produced more stable results, meaning that the mode point estimate was closer to the true reduction 0.7 more often than the not truncated data. In terms of uncertainty we can analyse the right part of Figure 2.7. This plot shows the width of every highest posterior density interval (HPDI: interval with limits the HPDLow95 and HPDHigh95 as seen in Section 2.3) and it is evident that the truncated data has more narrow intervals.

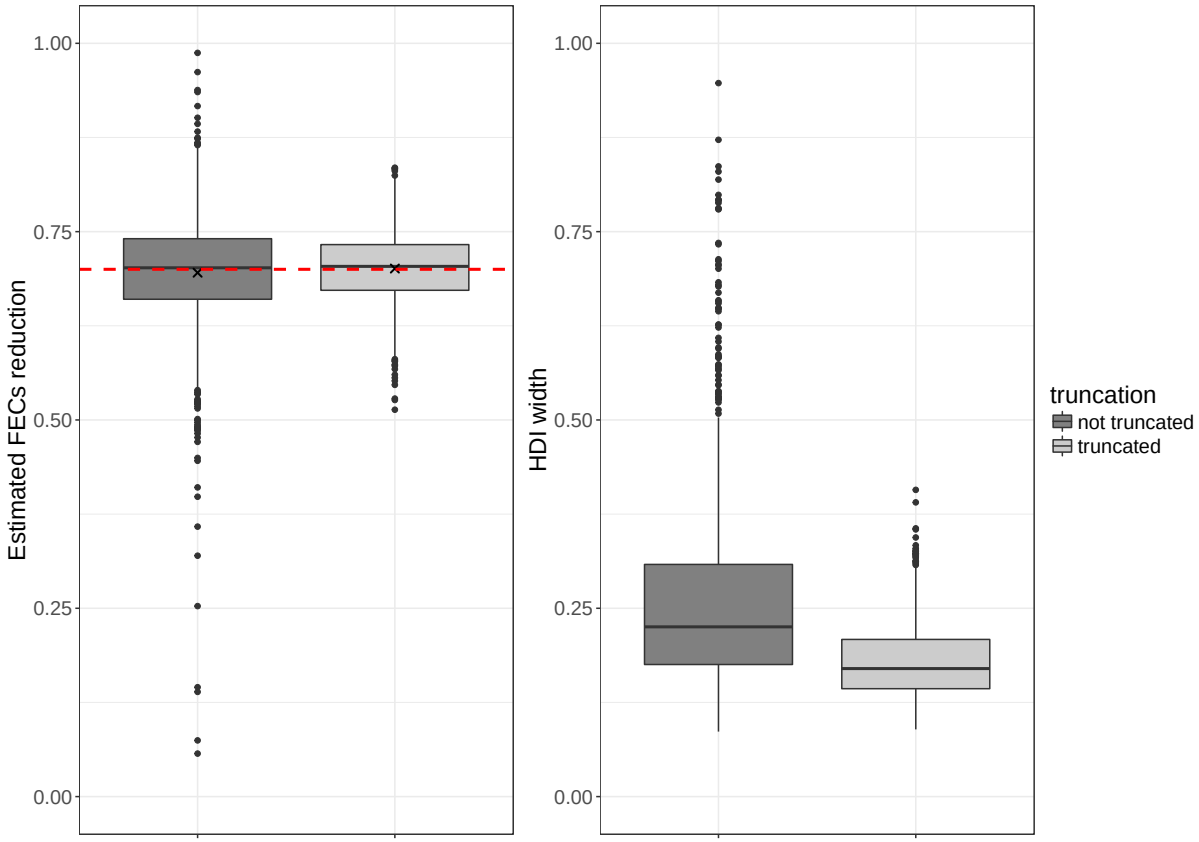

Figure 2.7: The left figure shows the 1000 mode point estimates for the mean FEC reduction together with the true reduction (red dashed line). The right figure reports the 1000 HPDI widths for each type of data.

These plots may lead a person to erroneously think that working with truncated data brings to more accurate results. The correct way of interpreting the output is to simply stress the difference between the datasets. If a study is conducted using only observations with initial epg larger than 200, then it is important to report this information. Generalizing the results from the truncated data to a non truncated case would underestimate the effect and lead to misleading results. In conclusion, truncated and non

truncated data should not be compared and, when conducting a study, all informations should be reported.

2.4.5 The dispersion and global mean parameters

In this section we work on the two hyperparameters that define the distribution of the individual-specific mean μ_i^b. As seen in section 2.2 the gamma distribution of μ_i^b has shape parameter k and rate parameter $\frac{k}{\mu}$, where k and μ are respectively the dispersion and global mean parameters. In the `eggCounts` package these hyperparameters were given specific distributions based on background knowledge of sheep and cattle. Because in this thesis we are studying the FECR for horses, appropriate considerations should be made. We start with the dispersion parameter k. Most papers about FEC and FECRT, including the paper from Torgerson *et al.* (2014) introducing the first version of the `eggCounts` package, agreee that within an observation group there is usually considerable overdispersion of FEC. Although the current version of the package does take overdispersion into account, we argue that the prior distribution Gamma(1, 0.7) assigned to the dispersion parameter k might not be the most appropriate. The reason is that the median of such distribution is 1, assigning a 50-50 chance for the data to be under- or over- dispersed, which does not reflect the hypothesis of overdispersed data just mentioned above. We hence propose a Gamma distribution with shape 1 and rate 0.35, which corresponds to an approximate 70% chance for overdispersion. Figure 2.8 illustrates the two distributions with respective probability for overdispersion.

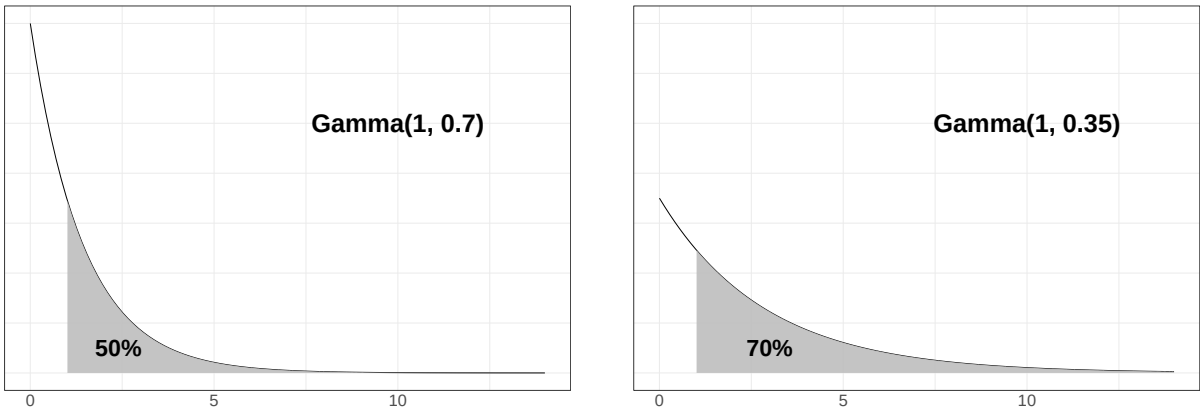

Figure 2.8: The prior distribution for the dispersion parameter. On the left side the current prior distribution as described in the eggCounts package and on the right side the proposed prior with a probability for overdispersion of approximately 70%.

Regarding the global mean μ, here as well we propose a different prior distribution based on the informations from Kaplan and Nielsen (2010), where 261 horses from 12 farms were studied. The article provides an estimate distribution for the adult equine population, recognizing that the actual distribution varies among farms and depends on multiple factors. The population is split into three groups: low, moderate and high contaminators (number of EPG). Table 2.1 gives the estimate numbers as reported in the article.

	Number of EPG (before treatment)	Percentage of adult population
Low	0–200 EPG	55
Moderate	200–500 EPG	18
High	>500 EPG	27

Table 2.1: Estimated population distribution of number of epg before treatment based on 261 horses from 12 farms (Kaplan and Nielsen (2010)).

We observe that most of the horses have an estimated number of EPG less than 200, while the prior distribution for the global mean defined in the `eggCounts` package assumes a prior global mean with much less weight for the interval 0–200 epg (see Figure 2.9). For this reason we propose a gamma distribution which reflects the distribution of the data of Table 2.1. Using the `multiroot` function included in the `rootSolve` in `R` (Soetaert (2009)) we compute the shape and rate values of the gamma distribution such that the probabilities for each category are as reported in Table 2.1. The distribution that satisfies the conditions is the Gamma(0.379, 0.00085) which is illustrated on the right side of Figure 2.9.

Although the prior distributions of both dispersion and global mean parameters are quite different to the original priors of the `eggCounts` package, we can observe on the left part of Figure 2.10 that the output of the reduction parameter is not sensible to the choice of these priors. What indeed is affected is the posterior distribution of the mean epg before treatment. On the right side of the same figure we can infact see that the original prior distributions from the package tend to "pull" the distribution of the mean epg before treatments towards higher values. This is a result of the larger weight given to higher epg values in the `eggCounts` prior. It is important to stress that the right side of Figure 2.10 is simply a sensitivity analysis and cannot state which is the "right" prior. Prior distributions are based on previous information and can vary from country to country or from

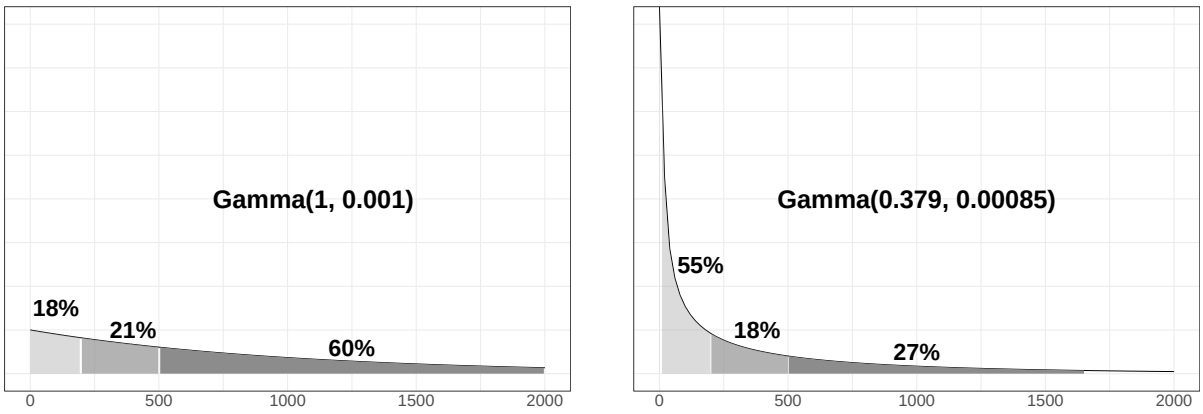

Figure 2.9: The prior distribution for the global mean parameter. On the left side the current prior distribution as described in the `eggCounts` package and on the right side the proposed prior based on the work of Kaplan and Nielsen (2010).

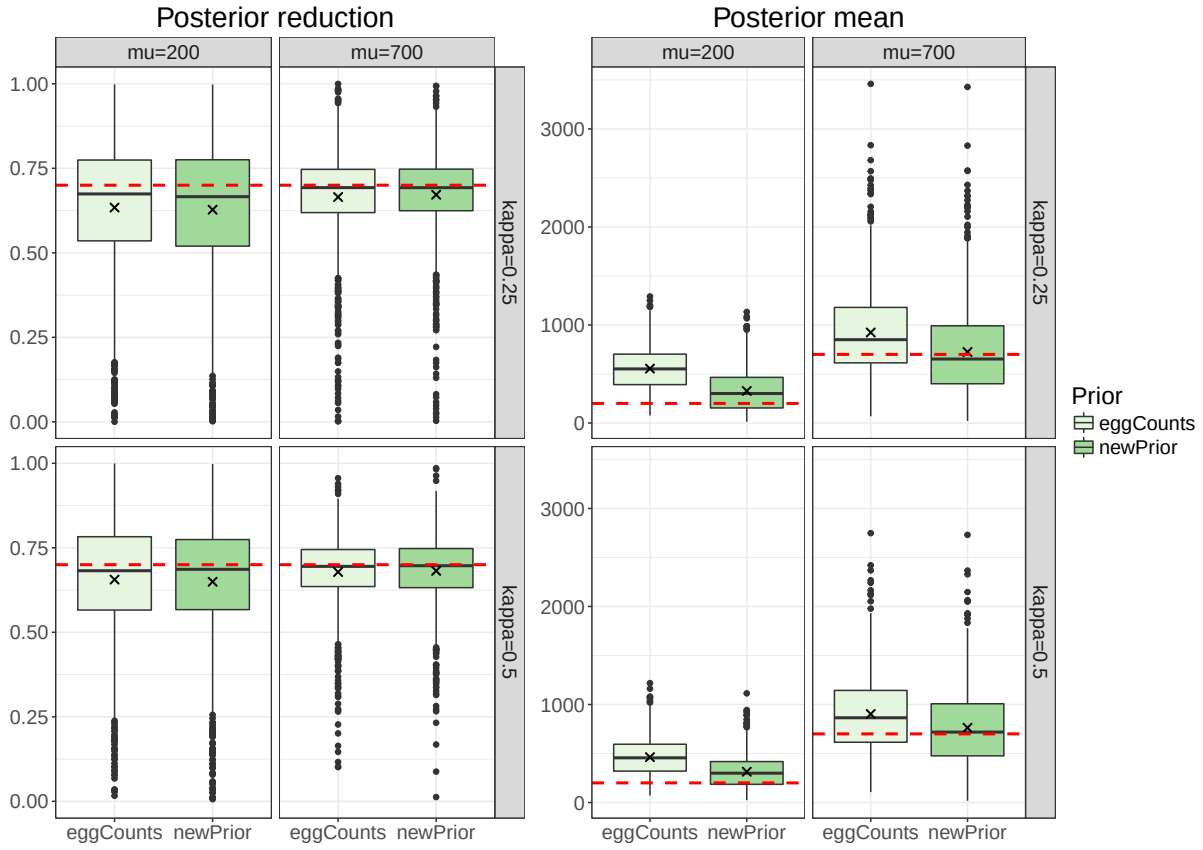

Figure 2.10: Boxplots of the estimated posterior reduction (left side) and posterior means (right side) using the `eggCounts` prior and the new prior Gamma(0.379,0.00085).

animal to animal.

One interesting aspect to point out is that until now the choice of these two priors was based on past papers and on the type of animal studied. We can further investigate the distributions of these priors by introduc-

ing the effect of truncated data described in Section 2.4.4. In particular we will study the distribution of the global mean μ when the counts before treatment are larger than 200 epg. As seen in Figure 2.9 the chosen prior for μ when conducting an equine analysis is a Gamma(0.379, 0.00085) which assignes a 55% probability to values between 0 and 200 epg. But what happens if we now exclude from the study animals with an initial epg lower than 200? We would expect the distribution of μ to shift towards higher numbers allowing for higher probabilities when the epg is higher than 200. How to compute the hyperparameters of this new "truncated" global mean distribution? We start by generating 100,000 numbers from the Gamma(0.379, 0.00085) and then only select the numbers higher than 200. This let us with a total of 44,856 numbers for which a gamma distribution is then fitted (see left side of Figure 2.11). This fitted Gamma distribution is then used as the new prior in the model. We can observe on the right side of Figure 2.11 that the shape of the curve has changed considerately compared to the Gamma(0.379, 0.00085) used previously but is now quite similar to the original Gamma(1, 0.001) from the `eggCounts` package. Although the probability for numbers between 200 and 500 has only slightly increased, the situation is almost inverted for datapoints lower than 200 and higher than 500.

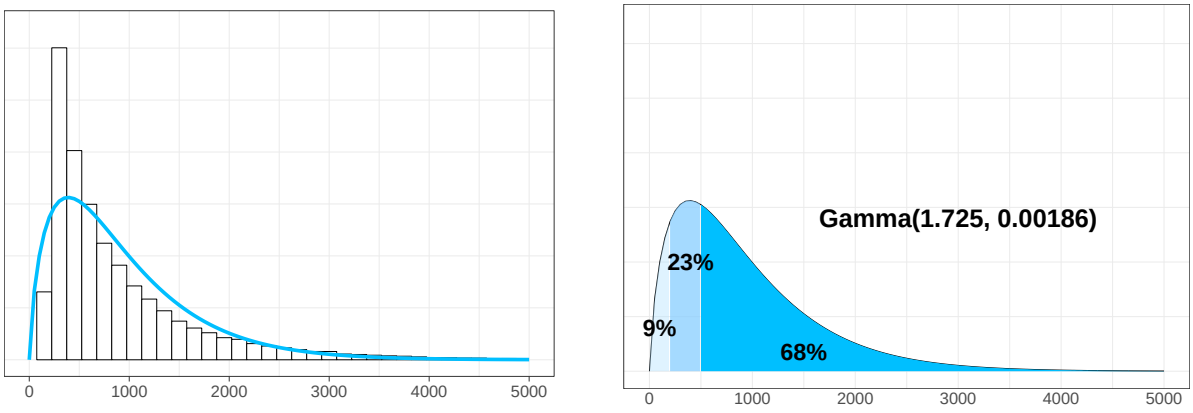

Figure 2.11: On the left side, the histogram of the 44,856 data points larger than 200. The shaded blue line is the fitted Gamma(1.725, 0.00186) distribution. On the right side the same Gamma(1.725, 0.00186) with highlighted probabilities for each group of epg (0-200, 200-500, >500).

To analyse the effect of this new fitted distribution we simulate 1000 datasets with truncated data (epg$\geq$200) varying μ=200epg, 700epg and k =0.25, 0.5 for both cases of 4 and 10 observations. The model is then fitted to each case using once the `eggCounts` prior and once the "truncated"

prior. The results are shown in Figure 2.12. Again we observe no significant difference in the output of the reduction parameter δ in any scenario. The widths of the HPDIs bring to the same conclusion of no difference (plot not shown).

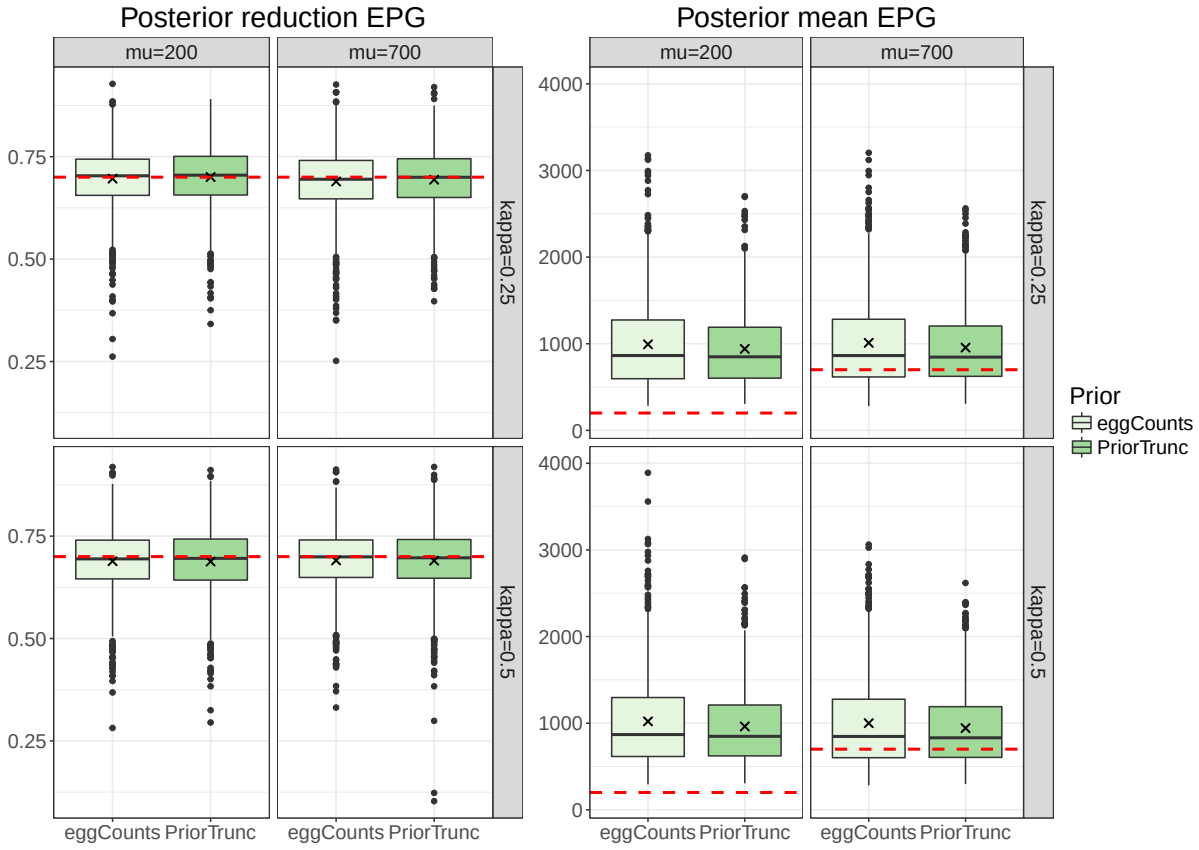

Figure 2.12: Boxplots of the estimated FECs reduction using the eggCounts and "truncated" priors for each of the 4 combinations of pre mean epg and dispersion value.

The fact that the parameter δ is not sensible to the priors of μ and k does not mean there is no difference in the output of other parameters. Infact, if we look at the plot of the mean epg before treatment (right part of Figure 2.12) it can be seen that here again there is some variation. Compared to Figure 2.10 the two distributions are now closer. The reason is that, as pointed out prevoiusly, the two prior distributions for the global mean μ are very similar.

2.4.6 Simple model

The `eggCounts` model used until now involves numerous parameters which despite increasing the robustness of the model they also raise complexity. In our context of very small samples a simpler model with less parameters

could be beneficial. In the previous section we have studied in detail the effect of the prior distributions of the global mean μ and dispersion parameter k on the output of the reduction parameter δ. In every case we have observed that the latter parameter was not affected. In this section we conduct a simulation study comparing the original `eggCounts` model with the simpler model where the two parameters μ and k are not taken into account. From a practical point of view this means we assume all animals come from a population with the same mean before tratment.

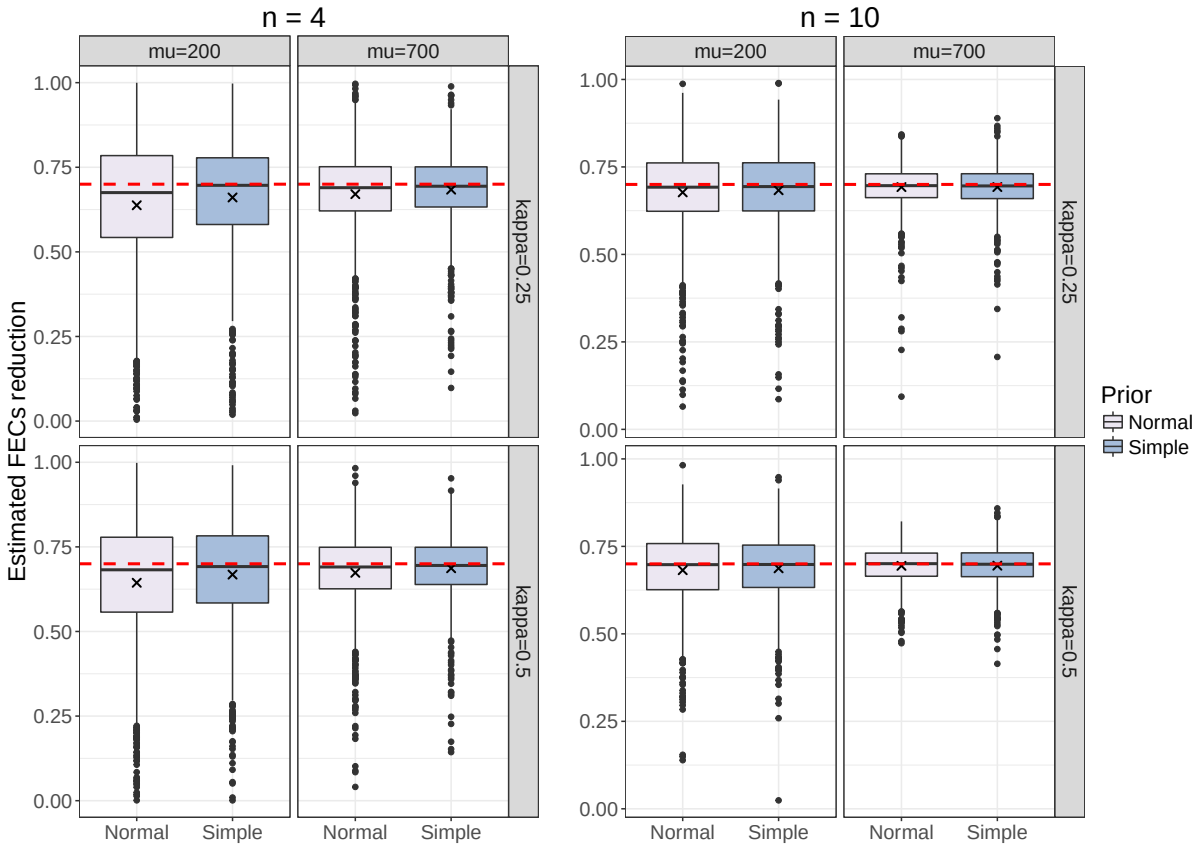

Figure 2.13: Boxplots of the estimated FECs reduction comparing the original model from the `eggCounts` package to the simpler model. The red dashed line is the true reduction used in the simulation.

The study, as before, is conducted by simulating 1000 datasets for 4 different scenarios varying the initial mean (200 epg and 700 epg) and the dispersion parameter (0.25 and 0.5). In the simple model a prior for the mean before treatment needs to be defined. In our analysis we choose to use the same prior Gamma(1, 0.001) used for the global mean in the original model. The results comparing the point estimates of the posterior distribution of the reduction parameter are shown in Figure 2.13. The

difference between the two models is quite evident in the case of $n = 4$ and specially when the number of epg is low. The simpler model seems to perform better in terms of both bias and variance. The conclusions when the sample size is larger ($n = 10$ in our case) are not the same. In this case we cannot observe a clear distinction between the two models, which is not necessarily a bad result. In fact, we can argue that the simple model performs as good as the more complex one, meaning that using this model which has less computation time and is easier to interpret we obtain similar results. We hence recommand practitioners to use the simple model when the number of animals studied is ≤ 10.

2.4.7 Interpreting the output

Now that we have listed some small sample considerations and finally run our model, all that remains is the interpretation of the output. In Section 1.2 we talked about the trichotomization of Coles *et al.* (1992), the different thresholds introduced for the different types of drug and the reason why this interpretation might not be the best choice. Here, we talk about posterior distributions and, based on the work of Wang *et al.* (2017b), give a more appropriate interpretation of the output.

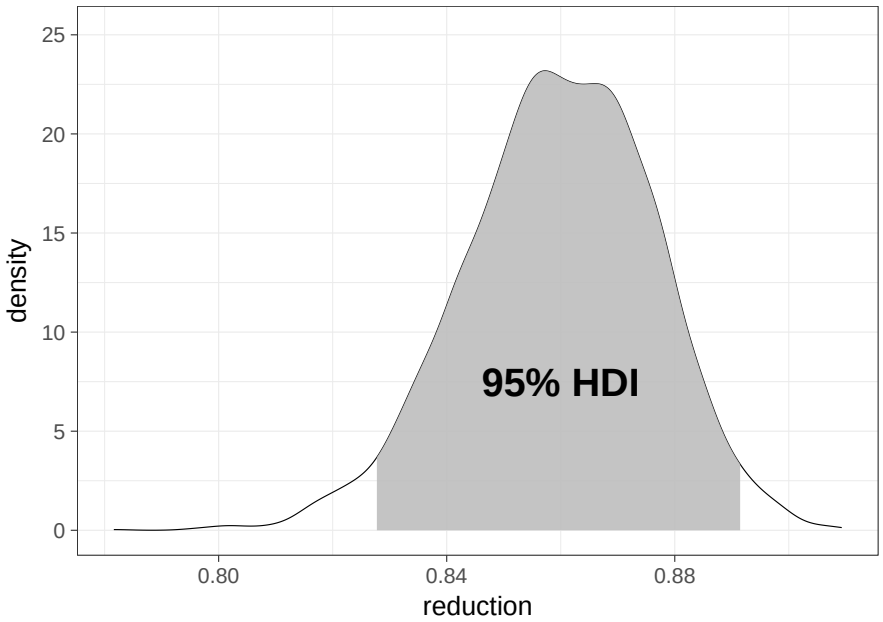

Figure 2.14: Density plot of the posterior ditribution for the reduction parameter with highlighted 95% HPDI (highest probability density interval).

Consider Figure 2.14 which shows the posterior distribution of the FEC

mean reduction after treatment with FECR mode of 0.86 and limits of the HPDI of 0.83 and 0.89. If we were to use the trichotomization from Coles *et al.* (1992) with thresholds 0.9 and 0.8 for FECR and LCL respectively, we would conclude there is suspected resistance because the LCL is larger that 0.8. Such conclusion can be misleading since we can observe that almost the whole distribution is less than 0.9. We hence propose to work with posterior distributions and compute the probability for the reduction to be lower than the thresholds defined in Coles *et al.* (1992) (0.9 for PYR and FBZ or 0.95 for IVM and MOX), which is equivalent to computing the probability that resistance is present. In our example the probability of resistance is 0.88.

For easier interpretation these probabilities can also be categorized. A possible classification could be:

- 0–20% very strong evidence of resistance,
- 20–40% strong evidence of resistance,
- 40–60% moderate evidence of resistance,
- 60–80% weak evidence of resistance,
- 80–100% no evidence of resistance.

In any case, we recommend to always report the value of the probability since practitioners may use different thresholds for their own classification.

This type of interpretation of the output is particularly beneficial when working with small samples since there is higher variability and the trichotomization from Coles *et al.* (1992) could lead to more frequent misclassifications.

Chapter 3

Case study: Anthelmintic Resistance in German horses

In this chapter we conduct a case study on anthelmintic resistance of German horses. The data, consisting of 14 horses in farm 1 and 6 horses in farm 2, was collected and analysed using the FECRT. The horses were treated with pyrantel (PYR) and the samples of epg were computed following the modified McMaster technique with an analytical sensitivity of 20. Here we re-analyze the data using the `eggCounts` package and including the considerations on small samples studied in the provious chapter. Furthermore, to help practitioners, we created a self explanatory flowchart (see Appendix A) that we follow step by step to illustrate its usage. The data from the two farms will be studied separately: we start with Farm 1.

3.1 Farm one: 14 observations

Following the flowchart, the first thing to check is if the number of animals tested is larger than 10. In this first case we have data from 14 horses hence the standard `eggCounts` will be used. The standard arguments in the function should be used except for the ones we will mention here below. Sequently it is required that the analytical sensitivity is ≤ 20. The practitioner reported a sensitivity of exactly 20 so we can specify this information in the function (`preCF=20`) and proceed to the next step. Notice that if a sensitivity larger than 20 would have been used, the model could have produced biased results, as the difference between an $f = 20$ and 30 is quite significant.

The following step is to specify if there is any prior information regarding the reduction parameter δ. This information could come from past

papers, general knowledge or from past results analysing the same sample. Such information should be chosen carefully since the output can be significantly affected. In our case the practitioner reported that "most of the times a reduction between 50% and 80% is observed". In our opinion, "most of the times" could translate to "7 times out of 10", but of course it is a very subjective statement and other analysts may interpret this information differently. Since we do have prior knowledge about the reduction and the extremes of a confidence interval are given, we use the `getPrior_delta` function (included in the `eggCounts` package) to compute the hyperparameters for the beta distribution. The function returns values of 5 and 2.38 for shape and rate respectively. This information is then included in the `fecr_stan` function by specifying
`deltaPrior = list(priorDist="beta", hyperpars=c(5,2.38))`.
The final code is shown below. The model can now finally be run.

```
fecr_stan(preFEC = FEC_Before,
          postFEC = FEC_After,
          rawCounts = FALSE,
          preCF = 20,
          paired = TRUE,
          zeroInflation=TRUE,
          deltaPrior = list(priorDist="beta",
                            hyperpars=c(5, 2.38)))
```

The results show a posterior mode of 0.75 with the limits of the 95%HDI of 0.68 and 0.8. Furthermore, as mentioned in the previous chapter, the advantage of working with Bayesian models is that we also obtain whole distributions we can use to draw conclusions. In the left part of Figure 3.1 such posterior distribution of the reduction parameter δ is plotted. We here enter the last step of the flowchart which is to interpret the output. Since the horses were treated with pyrantel, the cut-off value used to determine resistance is set to 90%. As can be seen in the figure, the whole distribution is below such threshold, indicating no evidence of resistance for Farm 1. Furthermore, the results from the FECRT produced a higher mode (0.83) and a much wider confidence interval: [0.38, 0.95] confirming the better performance of the hierarchical model described in Wang *et al.* (2017b).

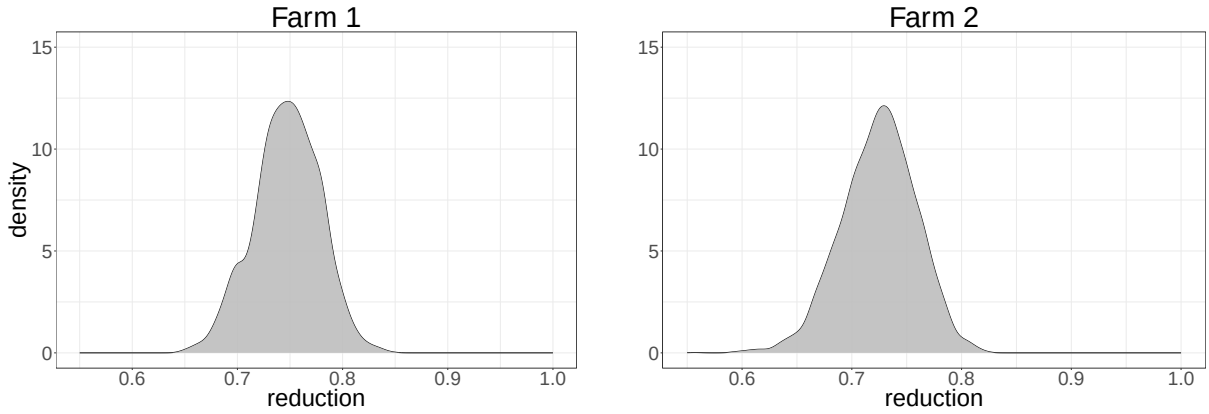

Figure 3.1: Posterior density distribution of the EPG reduction for each farm.

3.2 Farm two: 6 observations

The analysis for the data of the second farm is slightly different due to the small sample. The number of horses studied is 6 which, following the first step of the flowchart, suggests the simple model should be used. The new function `fecr_stanSimple` implements such model and, compared to the original `fecr_stan` model, it involves less parameters (no `paired` or `zeroInflation`). Next in the flowchart is the constraint regarding the analytical sensitivity. Here again we use an f of 20, meaning we can jump directly to the 3rd decision point concerning the prior distribution of δ. Since the data of the 2 farms come from the same practitioner, as before we use the information which assumes that 70% of the times the reduction will be between 50% and 80%. The `R` code in this case is reported below.

```
fecr_stanSimple(preFEC = FEC_Before,
                postFEC = FEC_After,
                rawCounts = FALSE,
                preCF = 20,
                deltaPrior = list(priorDist="beta",
                                  hyperpars=c(5, 2.38)))
```

The model is then finally run using the same hyperparameters for beta as above: 5 for shape and 2.38 for rate (obtained using the `getPrior_delta` function). The posterior distribution of δ is plotted on the right side of Figure 3.1. In this case the model produced a mode of 0.74 and a 95% HDI of [0.66, 0.79]. Since the whole distribution is below 90% the conclusion, as in farm 1, is that there is no evidence of resistance. In this case, using

the FECRT whould have produced a very similar mode (0.73) but a very wide confidence interval [0, 0.94] which would have brought to inconclusive results.

Chapter 4

Summary and conclusions

In this thesis we focused on the problem of very small sample size when analyzing the resistance of gastrointestinal nematodes in horses. Some of the numerous options available in the `eggCounts` package were used to partially reduce the high uncertainty resulting from small samples. Most of the work consisted in simulation studies comparing priors for different parameters.

We started by analysing datasets with different sensitivities and by comparing the techniques for faecal egg count reduction currently available on the market. The thesis then continued with more technical aspects regarding prior distributions for reduction, global mean and dispersion parameters. In particular, we proposed a simpler version of the `fecr_stan` function to use when analysing small samples. The `fecr_stanSimple` function is included in the `eggCounts` package version 2.0. All these considerations were then put into practice with a case study on the anthelmintic efficacy of German horses.

This thesis offers a detailed overview of the model of the `eggCounts` package when working with small samples. It provides an in depth study of the priors for equine analysis and an exhaustive explanation for practitioners of the several options included in the `eggCounts` package, including a guide and a flowchart which can be followed when conducting an anthelmintic efficacy study (see Appendix A and B).

We believe the analysed `eggCounts` model can be improved by introducing an animal specific reduction, meaning the hierarchical model would be further extended with two additional parameters. Moreover, we suggest to study in more detail the choice of the prior mean distribution of the new simple model suggested in Section 2.4.6 to verify if there is an influence of the type of animal analysed.

Appendix A

Flowchart

In this appendix the flowachart mentioned in Chapter 3 can be found. The steps are built for the analysis of anthelmintic resistance in horses but we believe they can be generalized to other animals such as sheep, cattle etc..

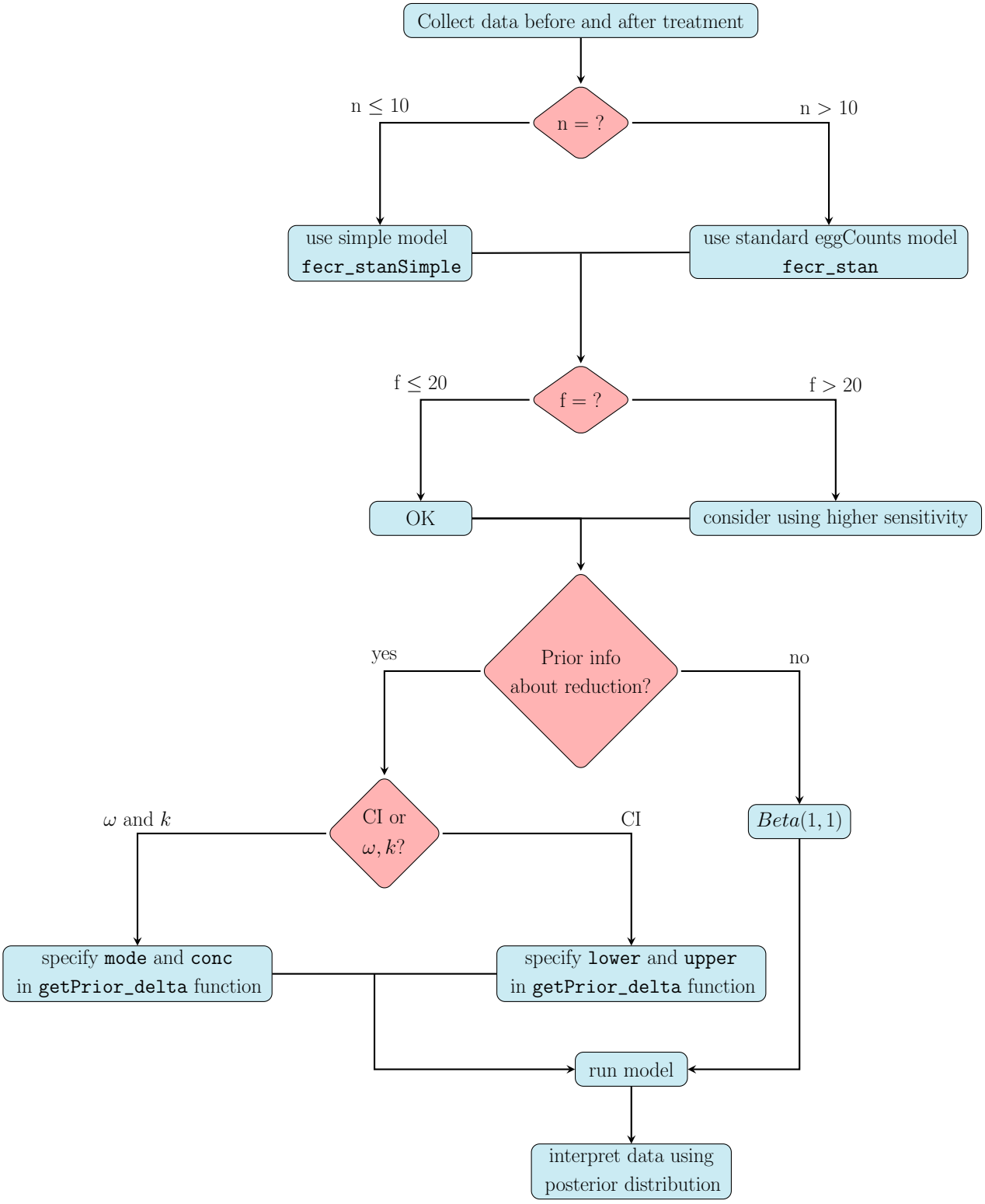

Figure A.1: Flowchart representing the steps to follow when performing an analysis of anthelmintic resistance in horses. At the start point it is assumed that the paired data (before/after treatment) has already been collected.

Appendix B

Guide for practitioners

When conducting an anthelmintic resistance drug test we recommand practitioners to use the hierarchical bayesian model described in the `eggCounts` package. This model provides more accurate estimates and lower uncertainty compared to the faecal egg counts reduction test from Coles *et al.* (1992) (see Wang *et al.* (2017b) for more detailed information). We here describe how an analysis in `R` is conducted, following the steps which are schematically represented in the flowchart. This guide is meant to help practitioners with very little `R` experience to analyse their data autonomously.

The steps to follow when analysing anthelmintic resistance drug tests are the following:

1. the data is first collected using the modified McMaster technique with the same analytical sensitivity, before and after treatment, of maximum 20 (eventually consider using the FLOTAC or mini-FLOTAC techniques which provide higher sensitivity)

2. the data should include at least two columns with number of egg counts before and after treatment (or treated and untreated groups)

3. if `RStudio` is being used, the data can be imported in `R` by clicking on "Import Dataset" located in the upper right corner

4. it should be checked that the data has been imported successfully (right number of rows and columns)

5. the analysis of the data can now start by loading the `eggCounts` package using the command `library(eggCounts)`

6. the first step requires the choice of the function: if n$\leq$ 10, `fecr_stanSimple` should be used, otherwise `fecr_stan`

7. if, i.e., the data is named `mydata`, then the options in both functions should be specified as follows:

   (a) `preFEC=mydata[,1]` if the data before treatment is located in the first column (otherwise replace 1 with 2)

   (b) `postFEC=mydata[,2]` if the data after treatment is located in the second column (otherwise replace 2 with 1)

   (c) `preCF=20` if a sensitivity of 20 was used (otherwise replace 20 with your factor)

   the name of the dataset can be seen in the upper right panel of `RStudio`

8. if any prior information regarding the reduction is available, the `getPrior_delta` function should be used to compute the hyperparameters, i.e. if 70% of the times the reduction is expected to be between 40% and 80%, the function should be specified as `getPrior_delta(lower=0.4, upper=0.8, p=0.7)`
   WARNING: for small samples, the output is heavily dependant on the specified prior distribution; the numbers should be chosen carefully

9. the `getPrior_delta` returns values of `alpha` and `beta` which can be included in the `fecr_stan` or `fecr_stanSimple` functions by specifying `deltaPrior=list(priorDist="beta", hyperpars=c(alpha,beta))`

10. after running the `fecr_stanSimple` (or `fecr_stan`) function, the output for the reduction can be found in the first line: "`fecr`"

11. we recommand to draw conclusions based on the probability of reduction, which can be computed using the function `fecr_probability`; the threshold should be specified based on the type of animal analysed and the type of drug used in the study

Appendix C

Shiny web interface

In the shiny web interface (Applied Statistics Group (2017)) we have implemented two options. The first one is related to the sample size, where the practitioner can specify the number of animal analysed and if such number is lower than ten then the new simple model described in Section 2.4.6 is used. The second implementation gives the possibility to specify a prior distribution for the reduction parameter by specifying either the lower and upper values of the confidence interval or the mode and concentration as described in Section 2.4.3.

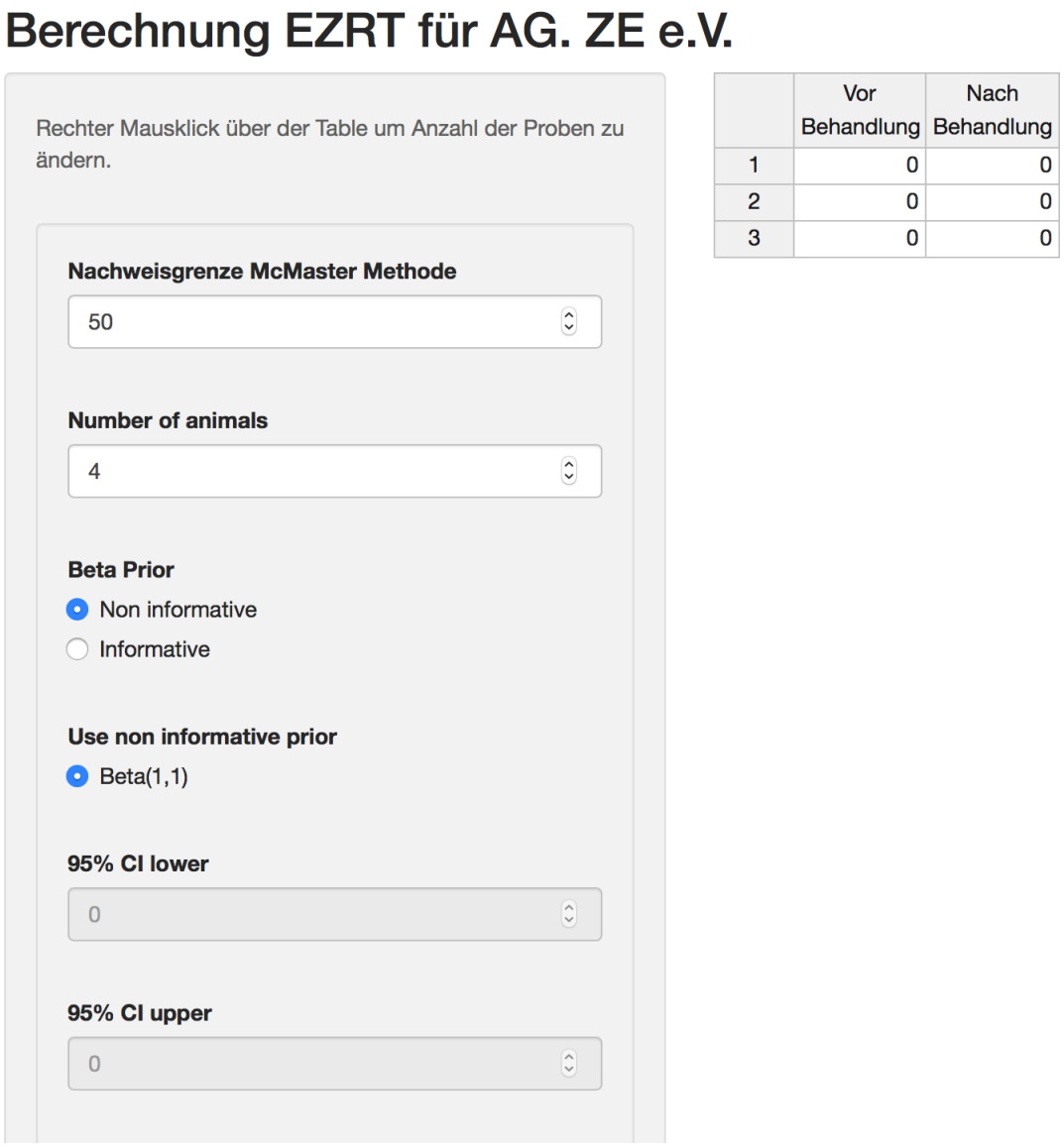

Figure C.1: The sample size option incorperated in the interface. In this figure the non informative beta prior is selected.

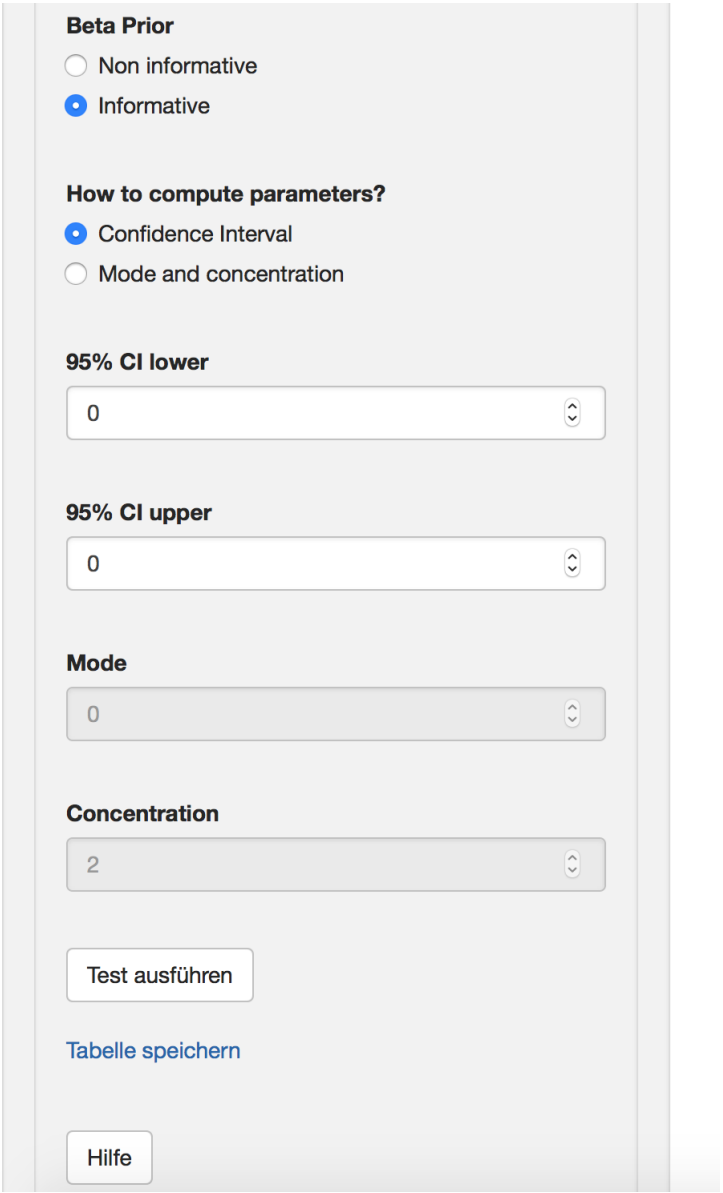

Figure C.2: The prior ditribution for beta is now set as informative and the user can specify either the lower and upper value of a confidence interval or the mode and concentration.

Bibliography

Applied Statistics Group (2017). Shiny app simpleegg. `shiny.math.uzh.ch/user/furrer/shinyas/shiny-simpleEgg/`.

Boehringer Ingelheim (2017). Endoparasites - life cycles. `http://www.merial.co.nz/Horse/diseaseinformation/Pages/en_life.aspx#n1`.

Coles, G. C., Bauer, C., Borgsteede, F. H. M., Geerts, S., Klei, T. R., Taylor, M. A., and Waller, P. J. (1992). World association for the advancement of veterinary parasitology methods for the detection of anthelmintic resistance in nematodes of veterinary importance. *Veterinary Parasitology*, **44**, 35 – 44.

Cringoli, G. (2006). FLOTAC, a novel apparatus for a multivalent faecal egg count technique. *Parassitologia*, **48**, 381 – 384.

Cringoli, G., Rinaldi, L., Albonico, M., Bergquist, R., and Utzinger, J. (2013). Geospatial (s) tools: Integration of advanced epidemiological sampling and novel diagnostics. *Geospatial Health*, **7**, 399 – 404.

Denwood, M. J., Reid, S. W. J., Love, S., Nielsen, M. K., Matthews, L., McKendrick, I. J., and Innocent, G. T. (2010). Comparison of three alternative methods for analysis of equine faecal egg count reduction test data. *Preventive Veterinary Medicine*, **93**, 316 – 323.

Fischer, J. K., Hinney, B., Denwood, M. J., Traversa, D., von Samson-Himmelstjerna, G., and Clausen, P.-H. (2015). Efficacy of selected anthelmintic drugs against cyathostomins in horses in the federal state of Brandenburg, Germany. *Parasitology Research*, **114**, 4441 – 4450.

Kaplan, R. M. and Nielsen, M. K. (2010). An evidence-based approach to equine parasite control: it ain's t the 60s anymore. *Equine Veterinary Education*, **22**, 306 – 316.

Kruschke, J. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan.* Academic Press.

Lester, H. E., Spanton, J., Stratford, C. H., Bartley, D. J., Morgan, E. R., Hodgkinson, J. E., Coumbe, K., Mair, T., Swan, B., Lemon, G., Cookson, R., and Matthews, J. B. (2013). Anthelmintic efficacy against cyathostomins in horses in Southern England. *Veterinary Parasitology*, **197**, 189 – 196.

McLeod, R. S. (1995). Costs of major parasites to the australian livestock industries. *International Journal for Parasitology*, **25**, 1363 – 1367.

Noel, M. L., Scare, J. A., Bellaw, J. L., and Nielsen, M. K. (2017). Accuracy and precision of mini-FLOTAC and McMaster techniques for determining equine strongyle egg counts. *Journal of Equine Veterinary Science*, **48**, 182 – 187.

Paul, M., Torgerson, P. R., Hoglund, J., and Furrer, R. (2014). Hierarchical modelling of faecal egg counts to assess anthelmintic efficacy.

Relf, V. E., Lester, H. E., Morgan, E. R., Hodgkinson, J. E., and Matthews, J. B. (2014). Anthelmintic efficacy on UK thoroughbred stud farms. *International Journal for Parasitology*, **44**, 507 – 514.

Sanna, G., Pipia, A. P., Tamponi, C., Manca, R., Varcasia, A., Traversa, D., and Scala, A. (2016). Anthelmintics efficacy against intestinal strongyles in horses of Sardinia, Italy. *Parasite Epidemiology and Control*, **1**, 15 – 19.

Soetaert, K. (2009). *Nonlinear Root finding, Equilibrium and Steady-state Analysis of Ordinary Differential Equations.*

Torgerson, P. R., Paul, M., and Furrer, R. (2014). Evaluating faecal egg count reduction using a specifically designed package "eggCounts" in R and a user friendly web interface. *International Journal for Parasitology*, **44**, 299 – 303.

Torgerson, P. R., Schnyder, M., and Hertzberg, H. (2005). Detection of anthelmintic resistance: A comparison of mathematical techniques. *Veterinary Parasitology*, **128**, 291 – 298.

van de Schoot, R., Broere, J. J., Perryck, K. H., Zondervan-Zwijnenburg, M., and van Loey, N. E. (2015). Analyzing small data sets using bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *European Journal of Psychotraumatology*, **6**, 25216.

Vidyashankar, A. N., Kaplan, R. M., and Chan, S. (2007). Statistical approach to measure the efficacy of anthelmintic treatment on horse farms. *Parasitology*, **134**, 2027 – 2039.

Wang, C., Furrer, R., and Paul, M. (2017a). *Hierarchical Modelling of Faecal Egg Counts*. `https://CRAN.R-project.org/package=eggCounts`.

Wang, C., Torgerson, P. R., Hoglund, J., and Furrer, R. (2017b). Zero-inflated hierarchical models for faecal egg counts to assess anthelmintic efficacy. *Veterinary Parasitology*, **235**, 20 – 28.