Statistical Analysis Plan v. 1
for B-Safe v. 0.1

Alexander Stemke (1), Dunfu Yang (1), Louise Whitehead (1,2)
reviewed by Oliver Sailer (1)
(1) Boehringer Ingelheim Pharma GmbH & Co. KG
(2) Biostatistics Research Group, Newcastle University, UK

2024-08-16

Contents

List of Tables

1 Purpose

The main purpose of the B-Safe Statistical Analysis Plan (SAP) is to provide a technical elaboration of the statistical methodologies and their computational implementation in the B-Safe application. B-Safe is an R-Shiny application for executing a Bayesian analysis via historical information borrowing. The B-Safe application was designed with safety analysis in mind. Therefore, the analysis deals with adverse events (AE). Specifically two types of adverse event metrics are considered in the B-Safe app, defined as follows: 1. Incidence proportion, the proportion of patients (out of all patients in a trial) having one or more specified adverse events during the follow up period. 2. Exposure-adjusted adverse event rate, defined as the number of adverse events per unit of time. Note that this is usually number of events per year, but could be any other unit of time (e.g., days, weeks, etc.). In this regard we only consider time to first event of a patient.

2 Introduction

This SAP focuses on the implementation of statistical and computational methods to calculate a meta-analytic-predictive (MAP) prior ([1]) and a robust extension ([2]) and perform inference on the derived posterior distribution by combining historical data with current data. For this, the R Bayesian Evidence Synthesis Tool package `RBesT` (v. 1.7-3) [3] is used.

For theoretical background, please refer to the documentation for `RBesT`. In this SAP we detail the default settings used in the app and our implementation of the RBesT functions.

The SAP is structured as follows. First, in Bayesian Computation (section 3) general concepts are briefly explained. The rest of the SAP follows the general structure of the B-Safe Shiny application. Data Preparation (section 4) covers the required format that the data should be. From MAP Prior (section 5) onward, there exist sub chapters corresponding to the analysis of incidence proportions or exposure adjusted event rates respectively. The general method or function call is explained at the beginning of a chapter with sub chapters focusing on the calculations and default settings.

The structure of this document is split into two parts. First, exploratory analysis, in which the user investigates a single variable (i.e. AE endpoint), using the app with headings corresponding to sections 2-5 (inclusive) in this SAP. Second, the Download (section 9), where all provided variables (i.e. AE endpoints), which are included in the uploaded dataset, will be analysed for each treatment arm.

3 Basics of Bayesian Computation

The computation of the meta analytic predictive prior involves two steps: computation of the posterior distribution $p(\theta \mid y)$ of a general parameter θ by use of the observed information y, the computation of the predictive distribution $p(\hat{y} \mid y, \theta)$ for some predicted future data $\hat{y}$.

For the computation to take place, the integral over a continuous function is evaluated by computing the value of the function of finite number of points. Desired accuracy is achieved by increasing the number of finite points.

For any parameter θ, the posterior distribution can be defined as:

$$p(\theta \mid y) = \frac{p(\theta)p(y \mid \theta)}{p(y)}.$$

$p(\theta)$ is the prior information, $p(y \mid \theta)$ the likelihood and $p(y)$ is the normalizing constant.

With Markov Chain Monte Carlo (MCMC) methods, `RBesT` obtains random samples of θ's from the desired distribution $p(\theta \mid y)$ and estimating the expectation of the posterior:

$$E(\theta \mid y) = \int \theta \cdot p(\theta \mid y) d\theta \approx \frac{1}{S} \sum_{s=1}^{S} \theta^{(s)}.$$

Other quantities are calculated in a similar way.

The accuracy of the simulation can be improved by obtaining more or larger randomly generated samples. Given each draw $\theta^{(s)}$, one can sample any predictive quantity $\tilde{y}$ from the predictive distribution $p(\tilde{y} \mid \theta)$. The set of simulated $\tilde{y}$'s from all the θ's characterize the posterior predictive distribution. The goal is to obtain a set of draws from the posterior distribution with enough draws so that the quantity of interest can be estimated with reasonable accuracy.

In MCMC, values of θ are drawn from an approximate distribution and then those draws are fitted to the target posterior distribution, $p(\theta \mid y)$. The sampling is done sequentially, making the distribution of the sampled draw depend on the last value drawn; hence they form a Markov chain. This is implemented in the B-Safe application using `RBesT` (s. [3]) which is an R-Software package.

`RBesT` uses `Stan` to create the MCMC samples. We have used the default MCMC settings `iter=6000, warmup=2000, thin=4, chains=4` which result in a posterior sample size of 4000.

4 Data Preparation

Data analysis is based on aggregated data per (sub-)study rather than patient level data. The summary statistics are provided using the following variables.:

Table 1: Data attributes

Variable	Data Type	Description
STUDYID	character/numeric	Study number or name
HIST	boolean	0 for current trial and 1 for historical trial
ARM	character/numeric	Defines the various treatment arms in the respective study
N	numeric	Total number of patients in the respective study
N WITH AE	numeric	Total number of patients who experienced at least one adverse event
SAF TOPIC	character	Safety endpoint variable to analyse , max 30 characters
TOT EXP	numeric	Total exposure time

The data could also include additional variables, such as `REGION` indicating the particular region where the study was conducted. When the data is available a treatment arm, `ARM`, has to be selected. All entries belonging to an `ARM` are used in the construction of the MAP prior and subsequent analysis of the corresponding safety endpoint `SAF_TOPIC`. Several rows in the dataset may belong to the same `ARM` and `SAF_TOPIC` (type of AE), but, for example be splitted into different regions. By selecting `pooling = TRUE` the number of patients, number of events and total exposure times over different regions will be displayed as the sum over the same `STUDY_ID`.

`TOT EXP` is the sum per `ARM` over all individual observed exposure times.

$$\texttt{TOT EXP} = \sum_{i=1}^{n} \mathrm{t}_i$$

where t_i represents the time until the first observed safety event.

Two types of analyses are provided to perform the safety assessment, "incidence proportion" and "exposure-adjusted AE rate". By default, all historical information (`HIST = 1`) for a specific `ARM` will be used to derive the MAP prior for a specified `ARM`.

For all computations a `seed` can be defined for reproducibility. If no seed is set, `as.numeric(Sys.time())` is used to set a seed.

5 MAP Prior

The Meta-Analytic Predictive (MAP) prior approach is a method of obtaining a prior distribution for a parameter of interest θ^* in the current study/clinical trial $p(\theta^* \mid Y_1, \ldots, Y_J)$. The MAP prior is derived using a Bayesian random effects meta-analysis model, i.e., the hierarchical model. The Bayesian meta-analysis accounts for unknown between-trial heterogeneity and the `RBesT` package samples the MAP Prior via Markov chain Monte Carlo (MCMC) algorithms as it is commonly not analytically tractable.

With trial and arm specific parameters $\theta^*, \theta_1, \ldots, \theta_J$, the simplest hierarchical model assumes the parameters underlying a particular arm are exchangeable across the trials and can be represented as:

$$\theta_j = \mu + \epsilon_j,\ \epsilon_j \sim N(0, \tau^2),\quad j = *, 1, \ldots, J.$$

The implementation of the MAP Prior approach requires the estimation of the overall mean, μ, and the between study variance τ^2 in a fully Bayesian way and therefore both μ and τ require specification of a prior distribution. Note that, in RBesT, μ is termed `beta`. In cases of low number of historical trials (studies), a sensitivity analysis of the posterior over a range of τ prior settings is recommended. Currently a `HalfNormal` prior distribution for τ is implemented.

The recommendation of Weber et. al (see [3]) is used in the BSafe application to quantify the degree of heterogeneity between trials, shown in the table below. σ refers to the standard deviation of the response variable, which is assumed known.

Table 2: Heterogeneity between trials

Heterogeneity	τ/σ
small	0.0625
moderate	0.125
substantial	0.25
large	0.5
very large	1

For both of the two types of analyses, implemented in the BSafe app (incidence proportions and exposure adjusted adverse event rates), the `RBesT::gMAP()` function is used to create the MAP prior.

To display summary statistics, `rstan` draws are directly accessed via `base` functions `mean(),sd(),quantile()` or directly by the `RBesT` provided function `RBesT::summary()`. Values are rounded to 4 digits and the 2.5%, 50% (Median) and 97.5% quantiles are displayed.

The posterior MCMC sample is fitted by `RBesT::automixfit()` using an expectation maximization approach. Any posterior sample can be described by a mixture distribution to any degree of precision. The number of mixture components is fixed to be 3 for computational reasons and has to be shown to sufficiently describe the posterior sample. The fitted distribution is used to display the MAP Prior.

5.0.1 Plots in general

For any plots `ggplot2` is used by sampling from their corresponding distributions. Values for the proportion are multiplied by 100 and displayed as %. Plots for the proportions range from $1 \cdot 10^{-5}$ to 1.

Plots for the exposure adjusted rate are always displayed on the log(rate) scale, i.e. as a Normal distribution. Plots for the rates range from the lowest 2% to the highest 98% quantile. The quantiles are in respect of the distribution to be displayed.

5.1 MAP Prior: Proportion

For incidence proportion, option `family=binomial` is set to assume a logit link function (`cbind(N WITH AE, N - N WITH AE)  1 | STUDYID`) in `RBesT::gMAP()`.

This results in the function call:

```
map_prop <- RBesT::gMAP(
  cbind(N_WITH_AE, N - N_WITH_AE) ~ 1 | STUDYID,
  data = input_data,
  tau.dist = tau_dist, # HalfNormal
  tau.prior = tau,
  beta.prior = 2,
  iter = 6000,
  warmup = 2000,
  thin = 4,
  chains = 4,
  family = "binomial"
)
```

The default choice for σ in the binomial case is 2 (as recommended in the `RBesT` documentation). This leads to a prior for beta N(0, σ^2) which is set in the gMAP function by specifying `beta.prior = 2`. This leads to the following choices for the between trial heterogeneity parameter τ, which has a prior distribution τ HalfNormal(0, τ^2) and is set within the `gMAP` function by specifying `tau.prior` = τ. Values for τ are chosen according to the user-specified level of between-trial heterogeneity, shown in the table below.

Table 3: Default τ hyperparameter values: proportion

Heterogeneity	τ
small	0.125
moderate	0.25
substantial	0.5
large	1
very large	2

For calculating the summary statistics (mean, standard error, median and the credible interval) of the MCMC samples of MAP prior for proportions the following code is used:

```
sample_prop <- rstan::extract(map_prop$fit)$theta_resp_pred
mean(sample_prop)
sd(sample_prop)
median(sample_prop)
quantile(sample_prop, probs = 0.025)
quantile(sample_prop, probs = 0.975)
```

Values for the proportion are multiplied by 100 and displayed as %.

The analytic approximation to the MCMC sample is fitted by `fit_prop <- RBesT::automixfit(map_prop, Nc = seq(3, 3))`

The forrest plot is included as property of the `RBesT::gMAP()` object `map_object`, called by `plot(map_object)$forest_model + bayesplot::legend_move("right")`.

5.2 MAP Prior: Exposure Adjusted Rate

For exposure-adjusted AE rate, the option `family=poisson` is set to assume a log link function. The hazard rate `TOT_EXP/N_WITH_AE` serves as parameter for the Poisson distribution. The response, `N_WITH_AE` is a vector of counts. The sum of total exposure times, `TOT_EXP`, can be specified by an offset, which will be linearly added to the linear predictor.

MCMC samples for the MAP prior for exposure adjusted adverse event rates are drawn using the following code:

```
map_rate <- RBesT::gMAP(N_WITH_AE ~ 1 + offset(log(TOT_EXP)) | STUDYID,
  data = input_data,
  tau.dist = tau_dist, # HalfNormal
  tau.prior = tau,
  beta.prior = 1,
  iter = 6000,
  warmup = 2000,
  thin = 4,
  chains = 4,
  family = "poisson"
)
```

The distribution of the log transformed hazard rates converge to a Normal distribution. The log hazards have a standard deviation of 1 (s. [4] p. 215). Therefore, σ is chosen to be 1 and `beta.prior = 1` which leads to the following default choices for τ, which has a prior distribution τ HalfNormal(0, τ^2) and is set within the `gMAP` function by specifying `tau.prior` = τ. Values for τ are chosen according to the user-specified level of between-trial heterogeneity, shown in the table below.

Table 4: Default τ hyperparameter values: exposure adjusted rate

Heterogeneity	τ
small	0.0625
moderate	0.125
substantial	0.25
large	0.5
very large	1

For exposure adjusted event rates, summary statistics (mean, standard error, median and the credible interval) of the MCMC samples of the MAP prior are calculated as follows:

```
sample_rate_log <- rstan::extract(map_rate$fit)$theta_pred
sample_rate_exp <- rstan::extract(map_rate$fit)$theta_resp_pred

# On the log hazard rate scale:
mean(sample_rate_log)
sd(sample_rate_log)
median(sample_rate_log)
quantile(sample_rate_log, probs = crilb)
quantile(sample_rate_log, probs = criub)

# On the hazard rate scale:
mean(sample_rate_exp)
sd(sample_rate_exp)
median(sample_rate_exp)
quantile(sample_rate_exp, probs = crilb)
quantile(sample_rate_exp, probs = criub)
```

An analytic approximation to the log hazard samples is then fitted:

```
theta_fit <- as.data.frame(map_rate$fit)
theta_pred <- theta_fit$theta_pred
fit_rate <- RBesT::automixfit(theta_pred, Nc = seq(3, 3))
```

5.3 Prior Effective Sample Size: Expected-Local Information Ratio (ELIR)

The amount of historical information contained in a prior can be expressed as the Effective Sample Size (ESS). The ESS indicates how much information the prior contains, in terms of how many experimental units the prior is roughly equivalent to. The ESS can be used to reduced the number of subjects in the control group of the current trial.

The Expected Local Information Ratio (ELIR) [5] method for calculating the ESS is used in B-Safe.

The ELIR method is commonly used as it is supposed to be predictively consistent. This implies that the expected posterior predictive sample size N is the sum of the prior expected sample size (ESS) and the sample size of the likelihood (observed data) of the current trial.

For both analyses (proportion and exposure adjusted rate) the ESS is calculated by `RBesT::ess()`.

For incidence rates, this is implemented by `RBesT::ess(fit_prop, method = "elir")`.

For exposure-adjusted event rates, this is implemented by `RBesT::ess(fit_rate, method = "elir", sigma = 1)`.

$\sigma = 1$ needs to be specified for the exposure adjusted rate analysis and in this case the ESS interpreted as the expected effective number of events (ENE) ([6]).

In some situations, unplausibly large values of ESS much greater than the total sample size of the historical data are reported. This is a known issue. In this case, please rerun the calculation with a different seed.

6 Robust MAP Prior

Historical data that are used in conjunction with a current trial run the risk of being in conflict. To mitigate this scenario, the MAP Prior is 'robustified'. In the robust MAP prior approach, a weakly informative (i.e., robust) component is added to the prior for the analysis of the current trial which will discount/down-weight the information contained in the historical trial:

$$\pi_{robustMAP} = (1 - w_r)\pi_{MAP} + w_r\pi_{non-informative}$$

where w_r indicates the weight on the robust component and π_{MAP} is the approximated MAP Prior and $\pi_{non-informative}$ is a vague prior. The choice of w_r is based on the user-specific degree of confidence and relevance placed on the historical data. The smaller the 'robustification' weight, the less discounting of the historical information and the more confidence in the information in the historical trial, and vice versa. When the current data and the historical data are similar, the posterior distribution of τ will have smaller variance indicating that the historical data is similar to the current data and stronger borrowing of historical information occurs.

The robustified MAP prior is obtained with the function `RBesT::robustify()`. This adds a fourth component to the fitted mixture distribution (`(fit_prop / fit_rate)`) with the chosen w_r. By default $w_r = 0.2$.

6.1 Robust MAP Prior: Proportion

For proportions, the robust MAP prior is implemented by `rob_map_prop <- RBesT::robustify(fit_prop, weight = rob_weight)` where `fit_prop` is the analytic approximation to the MAP prior for the proportion using a 3-component beta mixture distribution. `rob_weight` is the user-specified weight given to the non-informative component ($0 <$ `rob_weight` < 1).

For proportion analysis, the non-informative element of the MAP prior mean is by default set to 50% which represents no difference between the occurrence rates for one of the two outcomes, and a Beta(1,1) distribution is chosen as the weakly informative prior (as noted in the `RBesT` documentation, the uniform Beta(1,1) is appropriate in practical applications as the default robust prior) . The default prior weight on the weakly informative component has been set to 0.2 (20%).

Values for the proportion are multiplied by 100 and displayed as %.

6.2 Robust MAP Prior: Exposure Adjusted Rate

For exposure adjusted rates, this is implemented by `rob_map_rate <- RBesT::robustify(fit_rate, weight = rob_weight, mean = rob_mean, sigma = 1)`

where

`fit_rate` is the analytic approximation to the MAP prior for the log hazard rate using a 3-component normal mixture distribution `rob_weight` is the user-specified weight given to the non-informative component ($0 <$ `rob_weight` < 1). `rob_mean` is by default set to the mean of `fit_rate`. sigma = 1, as the Fisher information for 1 event = 1 leads to sigma = 1 ([6])

For the robustification of the exposure adjusted rate MAP prior, a weakly-informative 'non-exchangeability' parameter for the log-hazard in the current trial is added to the model with `sigma = 1`. If very heterogenous data is used to derive the MAP prior, the standard deviation of the informative part (MAP prior) might exceed sigma = 1. In that case, the intended uninformative robust part, would become informative. In that case we suggest to not borrow any information.

7 New Trial Data Analysis

For the New Trial Data Analysis, the robustified MAP prior is combined with information from a new study. `RBesT::postmix()` is used which calculates the posterior using conjugate updating of the prior with the new trial data. Displayed in the tables are the robustified MAP Prior, the likelihood information and the posterior.

The combined plots of robustified MAP prior, likelihood and posterior assists the user to detect a prior data conflict.

7.1 New Trial Data Analysis: Proportion

For incidence proportions, the posterior distribution for the new trial is calculated using `post_prop <- RBesT::postmix(rob_map_prop, n = n, r = r)`

where

`rob_map_prop` is the robust MAP prior, n is the sample size in the new trial and r is the number of observed events in the new trial.

The likelihood is displayed via `RBesT::mixbeta(lik = c(1, alpha = r, beta = n-r))`. Please note, that in the case of $r = n$, for the display of the likelihood as plot or as table, r is reduced by 1. Also, in the case of $r = 0$, r is set to 1. This does not impact the final posterior and was implemented to avoid plotting and calculation issues within `RBesT`.

7.2 New Trial Data Analysis: Exposure Adjusted Rate

For exposure adjusted rates, the posterior distribution for the new trial is calculated using `post_rate <- RBesT::postmix(rob_map_rate, m = log(new_hzrate), se = new_se)` where `rob_map_rate` is the robust map prior, `m` is `log(N_AE_new/TOT_EXP_new)`, `N_AE_new` is the number of patients with at least one AE in the new trial, `TOT_EXP_new` is the sum of their respective total exposure times and `new_se = sqrt(1/N_AE_new)`

The likelihood is displayed as `RBesT::mixnorm( lik = c(1, log(r/t), sqrt(1/r)))`.

8 Decision Making

Statistical inferences can be made based on the Likelihood, MAP Prior, Robust MAP Prior and the Posterior distribution. Note that the inference based on the likelihood is with respect to the observed data, in contrast to which based on the prior/posterior are with respect to the underlying parameter.

The user has the option to select values for an area of probability to be displayed with respect to the MAP Prior, robust MAP Prior, Likelihood and Posterior. If the area would be less than 1% or more then 99% of the probability mass, then the values will be fixed to 1% or respectively 99% to avoid display errors. Those fixed values don't influence any calculations, they are only for the display of the area under the curve and the statements according to those graphics.

The x values for `ggplot()` are calculated by:

```
length_disp <- 10000

if (select_analysis == "Incidence proportion") {
  x <- seq(0.009, 1, length = 10000)
} else if (select_analysis == "Exposure-adjusted AE rate") {
  a <- RBesT::qmix(param_approx, 0.009)
  b <- RBesT::qmix(param_approx, 0.991)
  x <- seq(a, b, length = length_disp)
  rm(a, b)
}
```

To calculate the y-values (probability and area under the curve), the fitted, robustified, updated posterior and "likelihood" distributions (expressed as `mix`) are used as argument in `RBesT::pmix(mix,...)`.

9 Download Results

One could directly jump to download summaries after providing data. All calculations are therefore, performed for each variable by the arms stated to be compared. Regarding MAP prior and posterior, the user has no option to change any default values (see section 9.3). Two for-loops run through each arm selection (`ARM` outer loop) and each variable (`SAF_TOP` inner loop). All calculations can be found in the script `ae_summary_table.R`.

The two groups will be analysed separately, i.e. all arms of the control group and all arms of the treatment group. For the naive estimation data will be pooled. For those naive estimation the proportions will be calculated by $\frac{\sum_{k=1}^{K} \texttt{N WITH AE}_k}{\sum_{k=1}^{K} \texttt{N}_k}$. And the naive exposure adjusted rate estimation is $\frac{\sum_{k=1}^{K} \texttt{N WITH AE}_k}{\sum_{k=1}^{K} \texttt{TOT EXP}_k}$, where K represents the number of groups/studies within the `ARM`.

Risk difference and risk ratio will be determined by sampling of each group a sample of size 10000 and taking the difference and the ratio and display the summary statistics.

One important note is that the calculations proceed if either historical information or current data is missing, in addition a warning text will be provided for which cases information was missing.

9.1 Missing information: Proportion

When no historical information is available a weak informative prior `RBesT::mixbeta(c(1, 1, 1))` is used. When no current data is available, the robustified MAP prior will be reported as posterior.

9.2 Missing information: Exposure Adjusted Rate

When no historical information is available, a weak informative prior will be used with the mean of the current data with SD=1 is used. When no current data is available, the robustified MAP prior will be reported as posterior.

9.3 Default Values

Default	Proportions	Rates
family	binomial	poisson
iter	6000	6000
warmup	2000	2000
thin	4	4
chains	4	4
gMAP: tau.dist	HalfNormal	HalfNormal
beta.prior	2	1
Large Heterogeneity	1	0.5
Robust Weight	0.2	0.2
Robust Mean	0.5	log(N WITH AE/TOT EXP)
Reference standard deviation	NA	1
No historical information	Beta(1,1)	Normal(log(N WITH AE NEW/TOT EXP NEW, 1)
No current data	Robustified MAP Prior as Posterior	Robustified MAP Prior as Posterior

References

[1] Neuenschwander, B., Capkun-Niggli, G., Branson, M. and Spiegelhalter, D. J. (2010). Summarizing historical information on controls in clinical trials. *Clinical Trials* **7** 5–18.

[2] Schmidli, H., Gsteiger, S., Roychoudhury, S., O'Hagan, A., Spiegelhalter, D. and Neuenschwander, B. (2014). Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* **70** 1023–32.

[3] Weber, S., Li, Y., Iii, J. W. S., Kakizume, T. and Schmidli, H. (2021). Applying meta-analytic-predictive priors with the *r* bayesian evidence synthesis tools. *Journal of Statistical Software* **100**.

[4] Aalen, B. and Gjessing. (2008). *Survival analysis & event history analysis: A process point of view.*

[5] Neuenschwander, B., Weber, S., Schmidli, H. and O'Hagan, A. (2020). Predictively consistent prior effective sample sizes. *Biometrics* **76** 578–87.

[6] Roychoudhury, S. and Neuenschwander, B. (2020). Bayesian leveraging of historical control data for a clinical trial with time-to-event endpoint. *Statistics in Medicine* **39** 984–95.