Untitled

18006513

2024-06-27

Now fit the models with half-cauchy prior for heterogeneity, and plot the forest plot.

```r
# Library relevant packages
library(brms)
library(metafor)
library(rstan)
library(kableExtra)
library(ggplot2)
library(bayesplot)
library(gridExtra)
library(ggplotify)
library(ggridges)
library(tidybayes)
library(dplyr)
library(glue)
library(stringr)
library(forcats)
library(posterior)

# Load data for each of the five outcomes
# with the first row containing variable names
infec.data <- read.csv("infection_data.csv", header = TRUE)
sym.infec.data <- read.csv("symptomatic_infection_data.csv", header = TRUE)
sev.infec.data <- read.csv("severe_infection_data.csv", header = TRUE)
hos.data <- read.csv("hospitalisation_data.csv", header = TRUE)
death.data <- read.csv("death_data.csv", header = TRUE)
```

```r
## Helper function to prepare data for Bayesian meta-analysis
# Calculate log relative risks, relative risks, variances, and SEs
prepare_data <- function(data){
  es <- escalc(measure = "RR",
               ai = n1, n1i = N1,
               ci = n2, n2i = N2,
               slab = author_id, data = data)
  es$RR <- exp(es$yi)
  es$sei <- sqrt(es$vi)

  # Remove the third and fourth columns corresponding to intervention indication variables
  es <- es[, -c(3, 4)]

  return(es)
```

```
}

# Prepare data for each outcome
infec.es <- prepare_data(infec.data)
sym.infec.es <- prepare_data(sym.infec.data)
sev.infec.es <- prepare_data(sev.infec.data)
hos.es <- prepare_data(hos.data)
death.es <- prepare_data(death.data)

## Define priors
# Model 2: Weakly Informative Priors with Half-Cauchy
priors_half_cauchy <- c(prior(normal(0, 1), class = Intercept),
                        prior(cauchy(0, 0.5), class = sd))
```

1.Infection

```
# Fit Bayesian meta-analysis model
infec.bayes <- brm(data = infec.es,
                   family = "gaussian",
                   yi | se(sei) ~ 1 + (1 | author_id),
                   prior = priors_half_cauchy,
                   iter = 2000, warmup = 1000, cores = 4, chains = 4,
                   control = list(adapt_delta = 0.99),
                   seed = 18006513)

# Summary of the Bayesian model
summary(infec.bayes)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = identity
## Formula: yi | se(sei) ~ 1 + (1 | author_id)
##    Data: infec.es (Number of observations: 15)
##   Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup draws = 4000
##
## Multilevel Hyperparameters:
## ~author_id (Number of levels: 15)
##               Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.43      0.10     0.28     0.66 1.01      616     1535
##
## Regression Coefficients:
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept    -0.37      0.13    -0.62    -0.13 1.00      647      964
##
## Further Distributional Parameters:
##       Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma     0.00      0.00     0.00     0.00   NA       NA       NA
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
# Rhat value for both parameters is 1, signifying convergence
pp_check(infec.bayes, ndraws = 100)
```
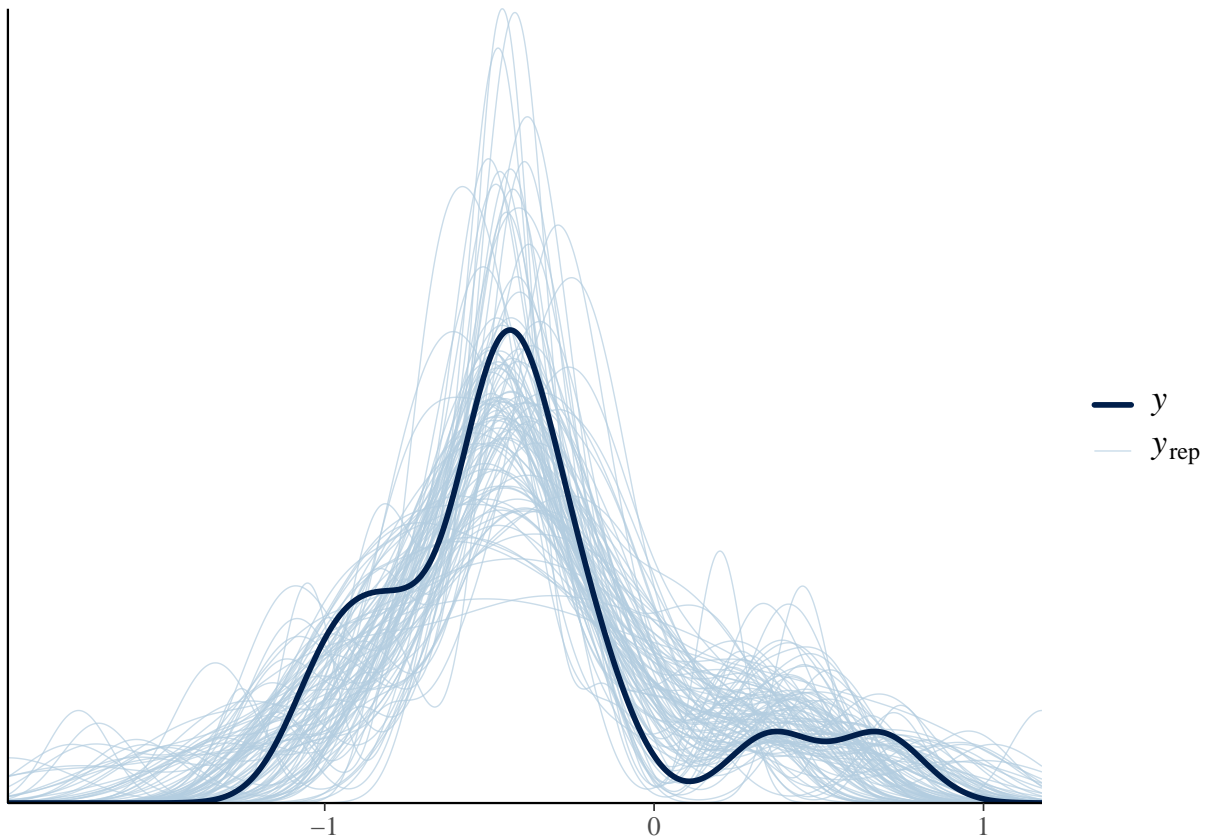


```
# Prepare Data for Forest Plot
# Spread draws for each study and compute b_Intercept
infec.draws <- spread_draws(infec.bayes, r_author_id[author_id,], b_Intercept) %>%
  mutate(b_Intercept = r_author_id + b_Intercept) %>%
  mutate(author_id = as.character(author_id))

# Pooled effect
infec.pooled.draws <- spread_draws(infec.bayes, b_Intercept) %>%
  mutate(author_id = "pooled effect")

# Combine individual study and pooled effect data frames
author_levels1 <- c(as.character(1:15), "pooled effect")
infec.forest <- bind_rows(infec.draws, infec.pooled.draws) %>%
  ungroup() %>%
  mutate(author_id = factor(author_id, levels = author_levels1))

# Calculate summary statistics
infec.summ <- group_by(infec.forest, author_id) %>%
  mean_qi(b_Intercept)

# Exponentiate to get risk ratios
infec.forest <- infec.forest %>%
```

```r
  mutate(b_Intercept = exp(b_Intercept))

infec.summ <- infec.summ %>%
  mutate(b_Intercept = exp(b_Intercept),
         .lower = exp(.lower),
         .upper = exp(.upper))

# Forest plot
ggplot(aes(x = b_Intercept, y = author_id), data = infec.forest) +

  # Add vertical lines for pooled effect and CI, and RR = 1
  geom_vline(xintercept = exp(fixef(infec.bayes)[1, 1]),
             color = "red", size = 1) +
  geom_vline(xintercept = exp(fixef(infec.bayes)[1, 3:4]),
             color = "red", linetype = 2) +
  geom_vline(xintercept = 1, color = "black",
             size = 1) +

  # Add densities
  geom_density_ridges(fill = "dodgerblue",
                      rel_min_height = 0.01,
                      col = NA, scale = 1,
                      alpha = 0.8) +
  geom_pointintervalh(data = infec.summ,
                      size = 1) +

  # Add text and labels
  geom_text(data = mutate_if(infec.summ, is.numeric, round, 2),
            aes(label = glue("{b_Intercept} [{.lower}, {.upper}]"), x = Inf), hjust = "inward") +
  labs(x = "Risk Ratio", y = "Author ID") +
  theme_minimal() +
  scale_y_discrete(limits = rev(levels(infec.summ$author_id)))
```
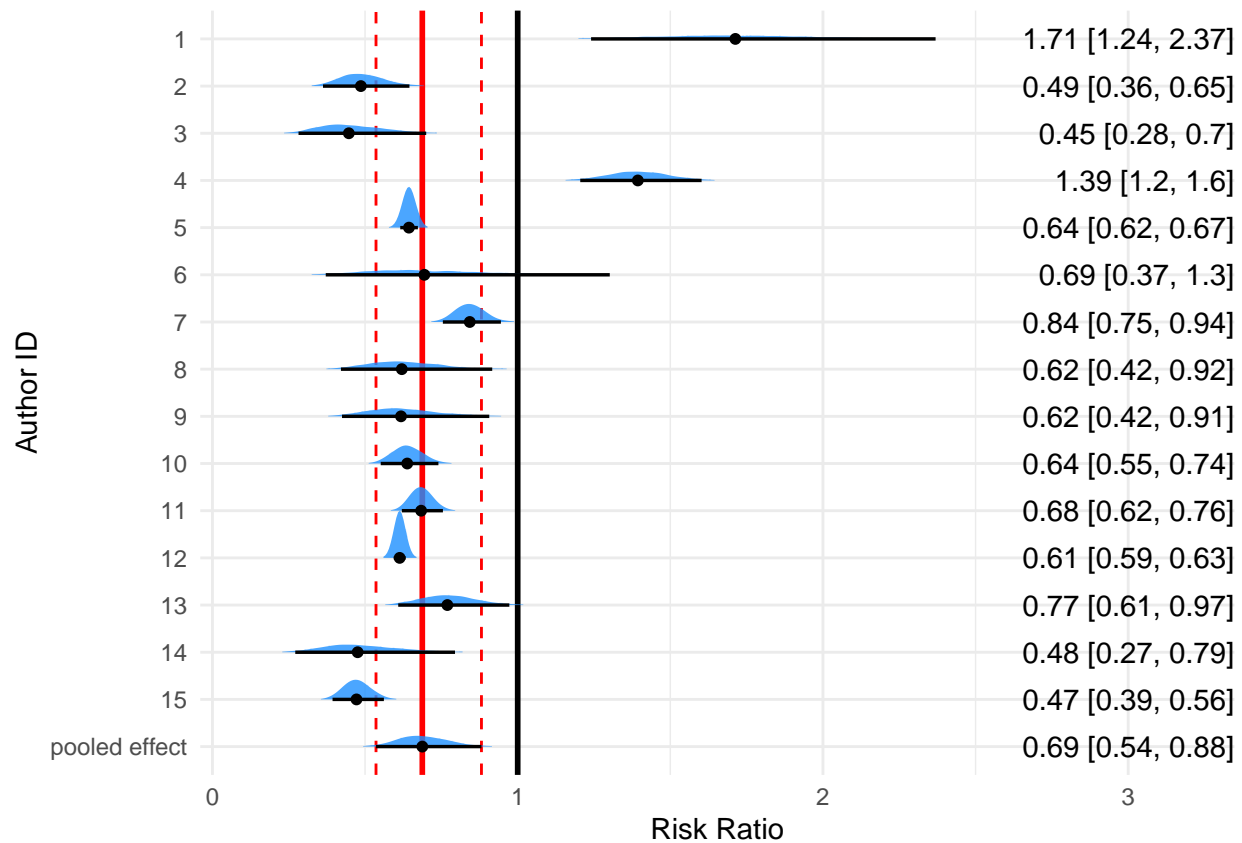
Author ID	Risk Ratio
1	1.71 [1.24, 2.37]
2	0.49 [0.36, 0.65]
3	0.45 [0.28, 0.7]
4	1.39 [1.2, 1.6]
5	0.64 [0.62, 0.67]
6	0.69 [0.37, 1.3]
7	0.84 [0.75, 0.94]
8	0.62 [0.42, 0.92]
9	0.62 [0.42, 0.91]
10	0.64 [0.55, 0.74]
11	0.68 [0.62, 0.76]
12	0.61 [0.59, 0.63]
13	0.77 [0.61, 0.97]
14	0.48 [0.27, 0.79]
15	0.47 [0.39, 0.56]
pooled effect	0.69 [0.54, 0.88]

2.Symptomatic Infection

```r
# Fit Bayesian meta-analysis model
sym.bayes <- brm(data = sym.infec.es,
                 family = "gaussian",
                 yi | se(sei) ~ 1 + (1 | author_id),
                 prior = priors_half_cauchy,
                 iter = 2000, warmup = 1000, cores = 4, chains = 4,
                 control = list(adapt_delta = 0.99),
                 seed = 18006513)

# Summary of the Bayesian model
summary(sym.bayes)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = identity
## Formula: yi | se(sei) ~ 1 + (1 | author_id)
##    Data: sym.infec.es (Number of observations: 5)
##   Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##         total post-warmup draws = 4000
##
## Multilevel Hyperparameters:
## ~author_id (Number of levels: 5)
```

```
##               Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.24      0.16     0.03     0.63 1.00      674      825
##
## Regression Coefficients:
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept    -0.36      0.14    -0.70    -0.10 1.00      949      960
##
## Further Distributional Parameters:
##        Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma      0.00      0.00     0.00     0.00   NA       NA       NA
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```r
# Rhat value for both parameters is 1, signifying convergence
pp_check(sym.bayes, ndraws = 100)
```
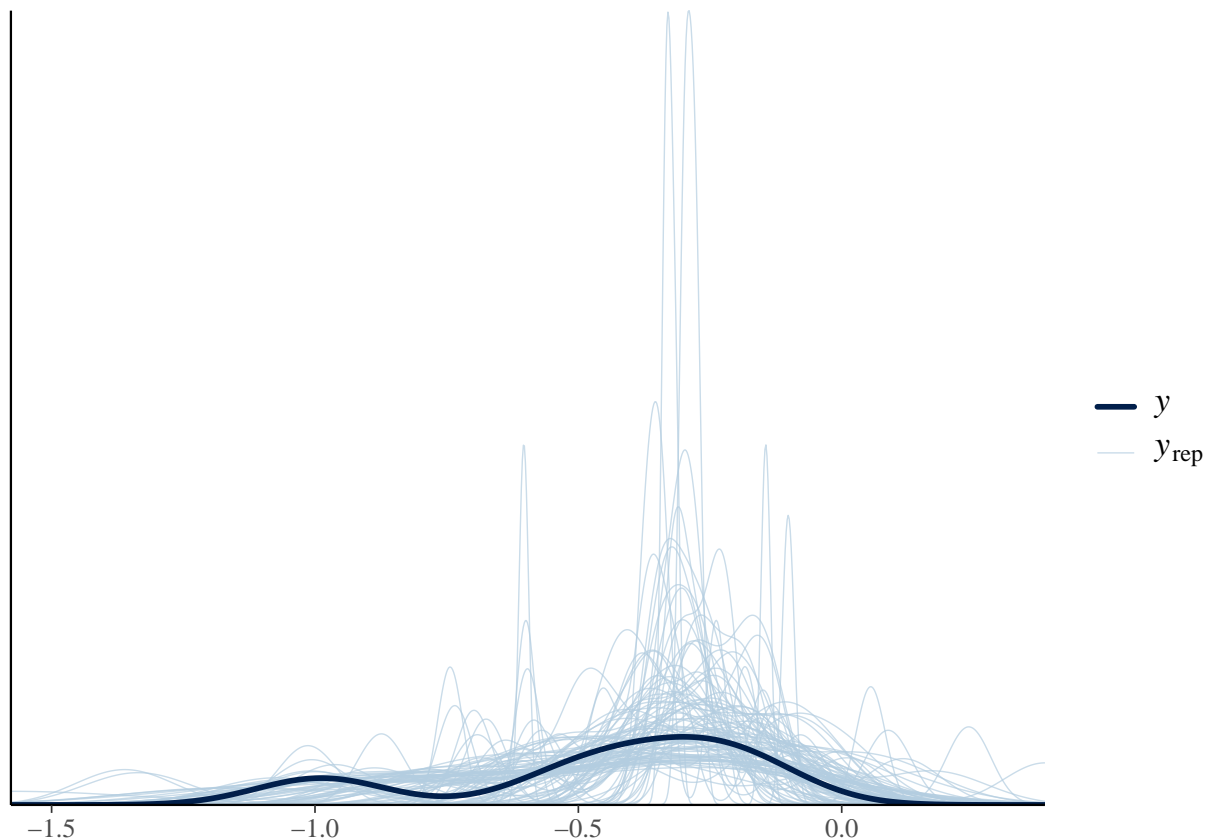


```r
# Prepare Data for Forest Plot
# Spread draws for each study and compute b_Intercept
sym.draws <- spread_draws(sym.bayes, r_author_id[author_id,], b_Intercept) %>%
  mutate(b_Intercept = r_author_id + b_Intercept) %>%
  mutate(author_id = as.character(author_id))

# Pooled effect
```

```r
sym.pooled.draws <- spread_draws(sym.bayes, b_Intercept) %>%
  mutate(author_id = "pooled effect")

# Combine individual study and pooled effect data frames
author_levels2<- c(as.character(1:5), "pooled effect")
sym.forest <- bind_rows(sym.draws, sym.pooled.draws) %>%
  ungroup() %>%
  mutate(author_id = factor(author_id, levels = author_levels2))

# Calculate summary statistics
sym.summ <- group_by(sym.forest, author_id) %>%
  mean_qi(b_Intercept)

# Exponentiate to get risk ratios
sym.forest <- sym.forest %>%
  mutate(b_Intercept = exp(b_Intercept))

sym.summ <- sym.summ %>%
  mutate(b_Intercept = exp(b_Intercept),
         .lower = exp(.lower),
         .upper = exp(.upper))

# Forest plot
ggplot(aes(x = b_Intercept, y = author_id), data = sym.forest) +

  # Add vertical lines for pooled effect and CI, and RR = 1
  geom_vline(xintercept = exp(fixef(sym.bayes)[1, 1]),
             color = "red", size = 1) +
  geom_vline(xintercept = exp(fixef(sym.bayes)[1, 3:4]),
             color = "red", linetype = 2) +
  geom_vline(xintercept = 1, color = "black",
             size = 1) +

  # Add densities
  geom_density_ridges(fill = "dodgerblue",
                      rel_min_height = 0.01,
                      col = NA, scale = 1,
                      alpha = 0.8) +
  geom_pointintervalh(data = sym.summ,
                      size = 1) +

  # Add text and labels
  geom_text(data = mutate_if(sym.summ, is.numeric, round, 2),
            aes(label = glue("{b_Intercept} [{.lower}, {.upper}]"), x = Inf), hjust = "inward") +
  labs(x = "Risk Ratio", y = "Author ID") +
  theme_minimal() +
  scale_y_discrete(limits = rev(levels(sym.summ$author_id)))
```
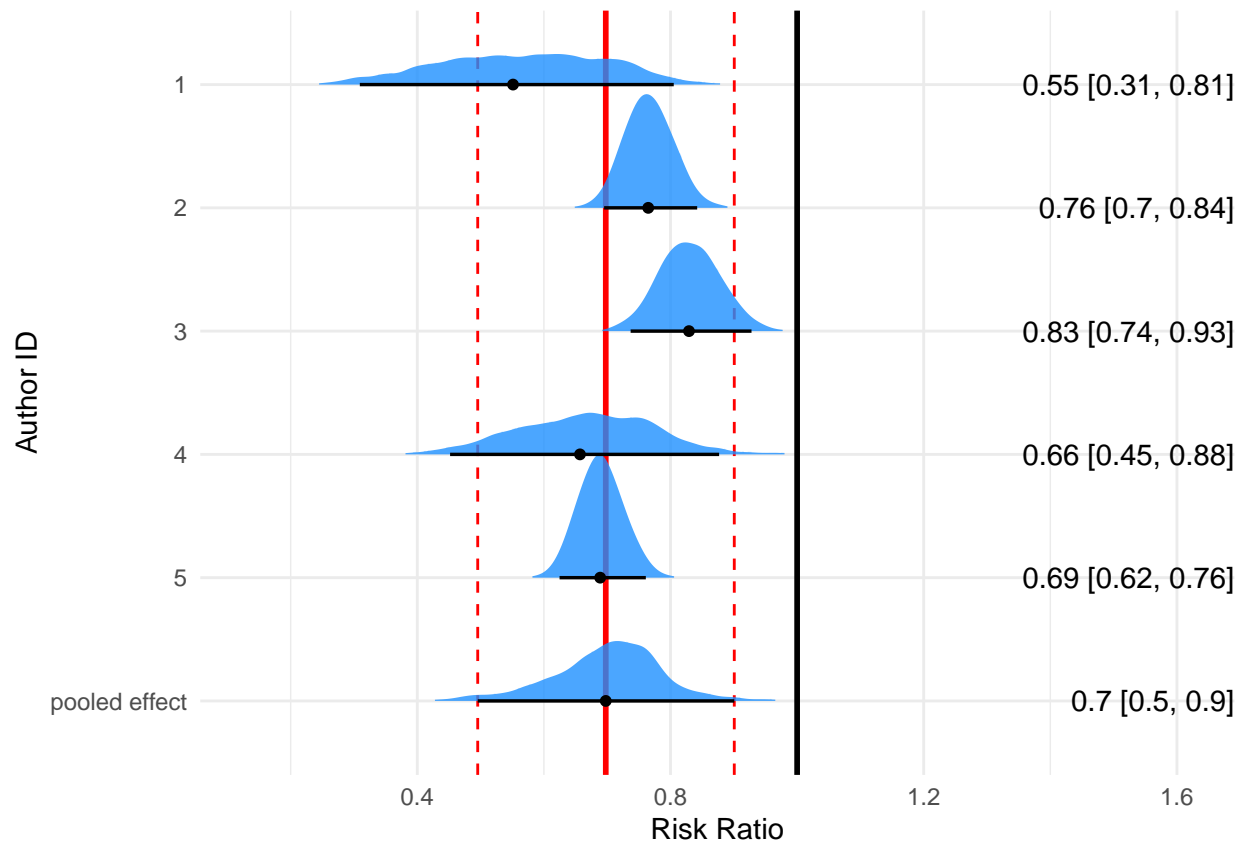
Author ID	Risk Ratio
1	0.55 [0.31, 0.81]
2	0.76 [0.7, 0.84]
3	0.83 [0.74, 0.93]
4	0.66 [0.45, 0.88]
5	0.69 [0.62, 0.76]
pooled effect	0.7 [0.5, 0.9]

3.Severe Infection

```
# Fit Bayesian meta-analysis model
sev.bayes <- brm(data = sev.infec.es,
                 family = "gaussian",
                 yi | se(sei) ~ 1 + (1 | author_id),
                 prior = priors_half_cauchy,
                 iter = 2000, warmup = 1000, cores = 4, chains = 4,
                 control = list(adapt_delta = 0.99),
                 seed = 18006513)

# Summary of the Bayesian model
summary(sev.bayes)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = identity
## Formula: yi | se(sei) ~ 1 + (1 | author_id)
##    Data: sev.infec.es (Number of observations: 6)
##   Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup draws = 4000
##
## Multilevel Hyperparameters:
## ~author_id (Number of levels: 6)
```

```
##                 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)       0.19      0.15     0.02     0.59 1.00      840     1201
##
## Regression Coefficients:
##            Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept     -0.44      0.13    -0.73    -0.21 1.00      867      646
##
## Further Distributional Parameters:
##         Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma       0.00      0.00     0.00     0.00   NA       NA       NA
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```r
# Rhat value for both parameters is 1, signifying convergence
pp_check(sev.bayes, ndraws = 100)
```
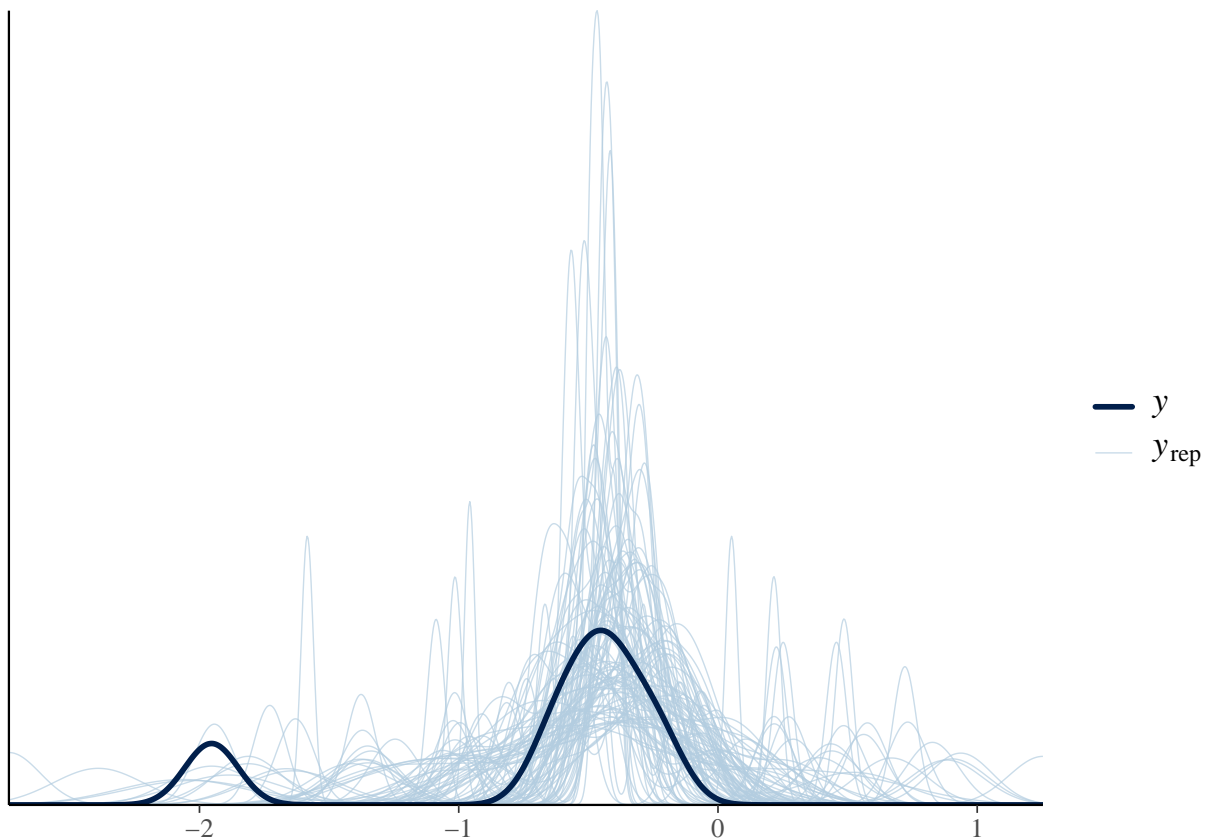


```r
# Prepare Data for Forest Plot
# Spread draws for each study and compute b_Intercept
sev.draws <- spread_draws(sev.bayes, r_author_id[author_id,], b_Intercept) %>%
  mutate(b_Intercept = r_author_id + b_Intercept) %>%
  mutate(author_id = as.character(author_id))

# Pooled effect
```

9

```r
sev.pooled.draws <- spread_draws(sev.bayes, b_Intercept) %>%
  mutate(author_id = "pooled effect")

# Combine individual study and pooled effect data frames
author_levels3<- c(as.character(1:6), "pooled effect")
sev.forest <- bind_rows(sev.draws, sev.pooled.draws) %>%
  ungroup() %>%
  mutate(author_id = factor(author_id, levels = author_levels3))

# Calculate summary statistics
sev.summ <- group_by(sev.forest, author_id) %>%
  mean_qi(b_Intercept)

# Exponentiate to get risk ratios
sev.forest <- sev.forest %>%
  mutate(b_Intercept = exp(b_Intercept))

sev.summ <- sev.summ %>%
  mutate(b_Intercept = exp(b_Intercept),
         .lower = exp(.lower),
         .upper = exp(.upper))

# Forest plot
ggplot(aes(x = b_Intercept, y = author_id), data = sev.forest) +

  # Add vertical lines for pooled effect and CI, and RR = 1
  geom_vline(xintercept = exp(fixef(sev.bayes)[1, 1]),
             color = "red", size = 1) +
  geom_vline(xintercept = exp(fixef(sev.bayes)[1, 3:4]),
             color = "red", linetype = 2) +
  geom_vline(xintercept = 1, color = "black",
             size = 1) +

  # Add densities
  geom_density_ridges(fill = "dodgerblue",
                      rel_min_height = 0.01,
                      col = NA, scale = 1,
                      alpha = 0.8) +
  geom_pointintervalh(data = sev.summ,
                      size = 1) +

  # Add text and labels
  geom_text(data = mutate_if(sev.summ, is.numeric, round, 2),
            aes(label = glue("{b_Intercept} [{.lower}, {.upper}]"), x = Inf), hjust = "inward") +
  labs(x = "Risk Ratio", y = "Author ID") +
  theme_minimal() +
  scale_y_discrete(limits = rev(levels(sev.summ$author_id)))
```
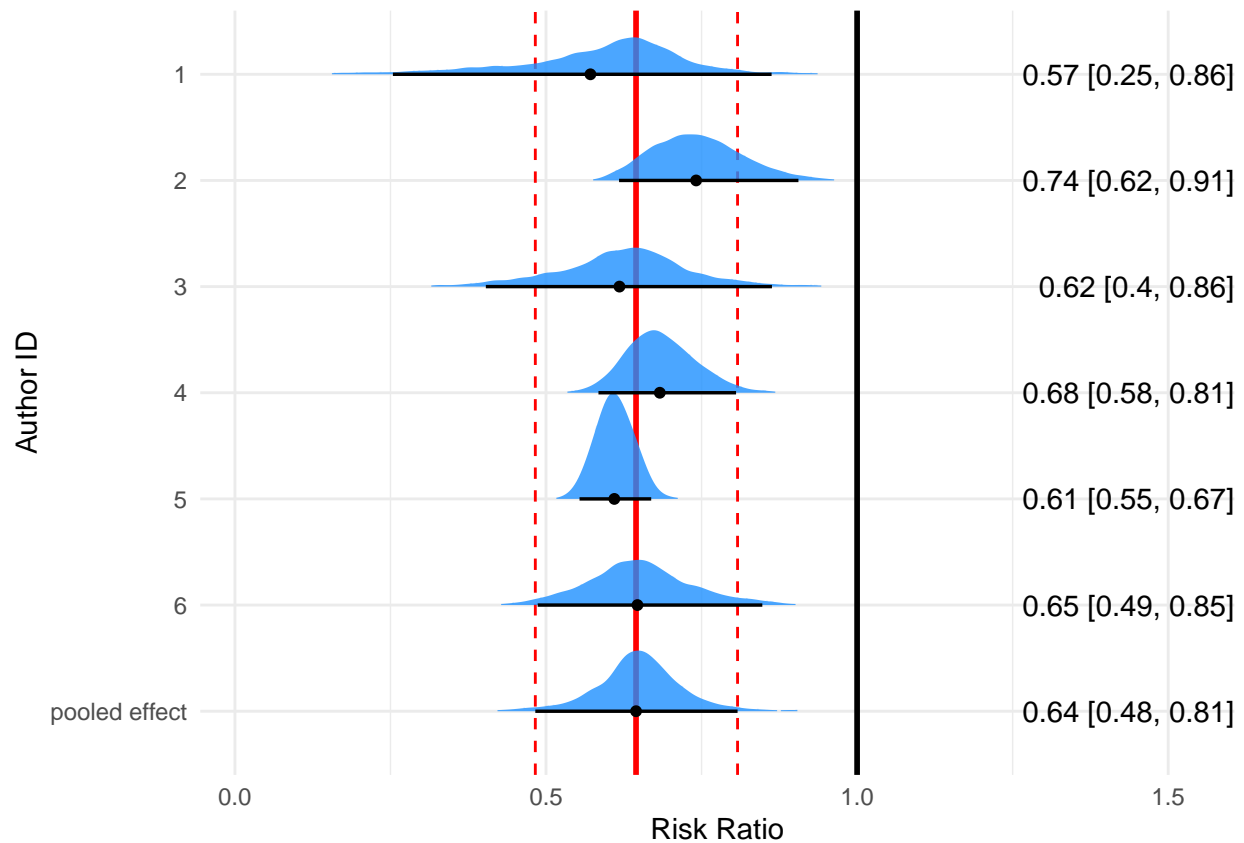
1	0.57 [0.25, 0.86]
2	0.74 [0.62, 0.91]
3	0.62 [0.4, 0.86]
4	0.68 [0.58, 0.81]
5	0.61 [0.55, 0.67]
6	0.65 [0.49, 0.85]
pooled effect	0.64 [0.48, 0.81]

4.Hospitalisation

```r
# Fit Bayesian meta-analysis model
hos.bayes <- brm(data = hos.es,
                 family = "gaussian",
                 yi | se(sei) ~ 1 + (1 | author_id),
                 prior = priors_half_cauchy,
                 iter = 2000, warmup = 1000, cores = 4, chains = 4,
                 control = list(adapt_delta = 0.99),
                 seed = 18006513)

# Summary of the Bayesian model
summary(hos.bayes)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = identity
## Formula: yi | se(sei) ~ 1 + (1 | author_id)
##    Data: hos.es (Number of observations: 4)
##   Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup draws = 4000
##
## Multilevel Hyperparameters:
## ~author_id (Number of levels: 4)
```

```
##                  Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)        0.19      0.20     0.01     0.73 1.00      709     1560
##
## Regression Coefficients:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept      -0.50      0.15    -0.90    -0.26 1.00      754      570
##
## Further Distributional Parameters:
##         Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma       0.00      0.00     0.00     0.00   NA       NA       NA
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```r
# Rhat value for both parameters is 1, signifying convergence
pp_check(hos.bayes, ndraws = 100)
```
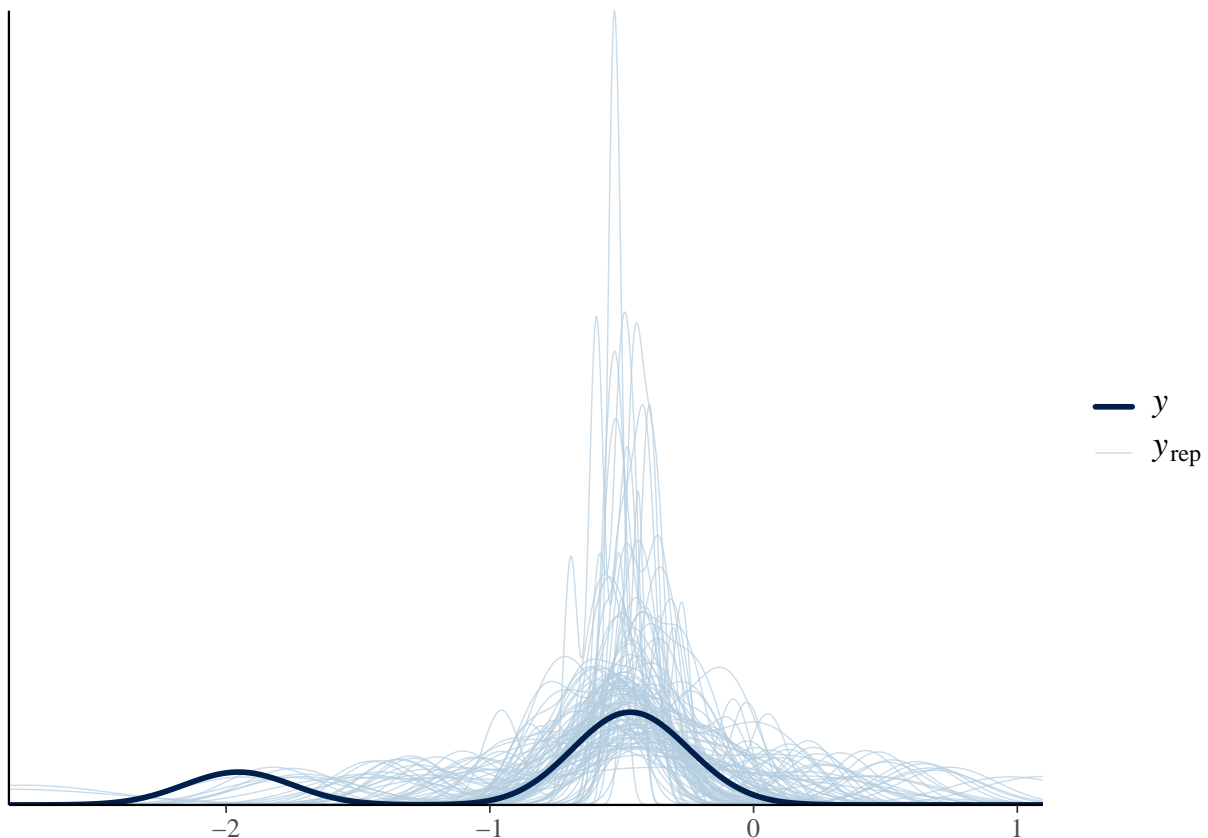


```r
# Prepare Data for Forest Plot
# Spread draws for each study and compute b_Intercept
hos.draws <- spread_draws(hos.bayes, r_author_id[author_id,], b_Intercept) %>%
  mutate(b_Intercept = r_author_id + b_Intercept) %>%
  mutate(author_id = as.character(author_id))

# Pooled effect
```

12

```r
hos.pooled.draws <- spread_draws(hos.bayes, b_Intercept) %>%
  mutate(author_id = "pooled effect")

# Combine individual study and pooled effect data frames
author_levels4<- c(as.character(1:4), "pooled effect")
hos.forest <- bind_rows(hos.draws, hos.pooled.draws) %>%
  ungroup() %>%
  mutate(author_id = factor(author_id, levels = author_levels4))

# Calculate summary statistics
hos.summ <- group_by(hos.forest, author_id) %>%
  mean_qi(b_Intercept)

# Exponentiate to get risk ratios
hos.forest <- hos.forest %>%
  mutate(b_Intercept = exp(b_Intercept))

hos.summ <- hos.summ %>%
  mutate(b_Intercept = exp(b_Intercept),
         .lower = exp(.lower),
         .upper = exp(.upper))

# Forest plot
ggplot(aes(x = b_Intercept, y = author_id), data = hos.forest) +

  # Add vertical lines for pooled effect and CI, and RR = 1
  geom_vline(xintercept = exp(fixef(hos.bayes)[1, 1]),
             color = "red", size = 1) +
  geom_vline(xintercept = exp(fixef(hos.bayes)[1, 3:4]),
             color = "red", linetype = 2) +
  geom_vline(xintercept = 1, color = "black",
             size = 1) +

  # Add densities
  geom_density_ridges(fill = "dodgerblue",
                      rel_min_height = 0.01,
                      col = NA, scale = 1,
                      alpha = 0.8) +
  geom_pointintervalh(data = hos.summ,
                      size = 1) +

  # Add text and labels
  geom_text(data = mutate_if(hos.summ, is.numeric, round, 2),
            aes(label = glue("{b_Intercept} [{.lower}, {.upper}]"), x = Inf), hjust = "inward") +
  labs(x = "Risk Ratio", y = "Author ID") +
  theme_minimal() +
  scale_y_discrete(limits = rev(levels(hos.summ$author_id)))
```
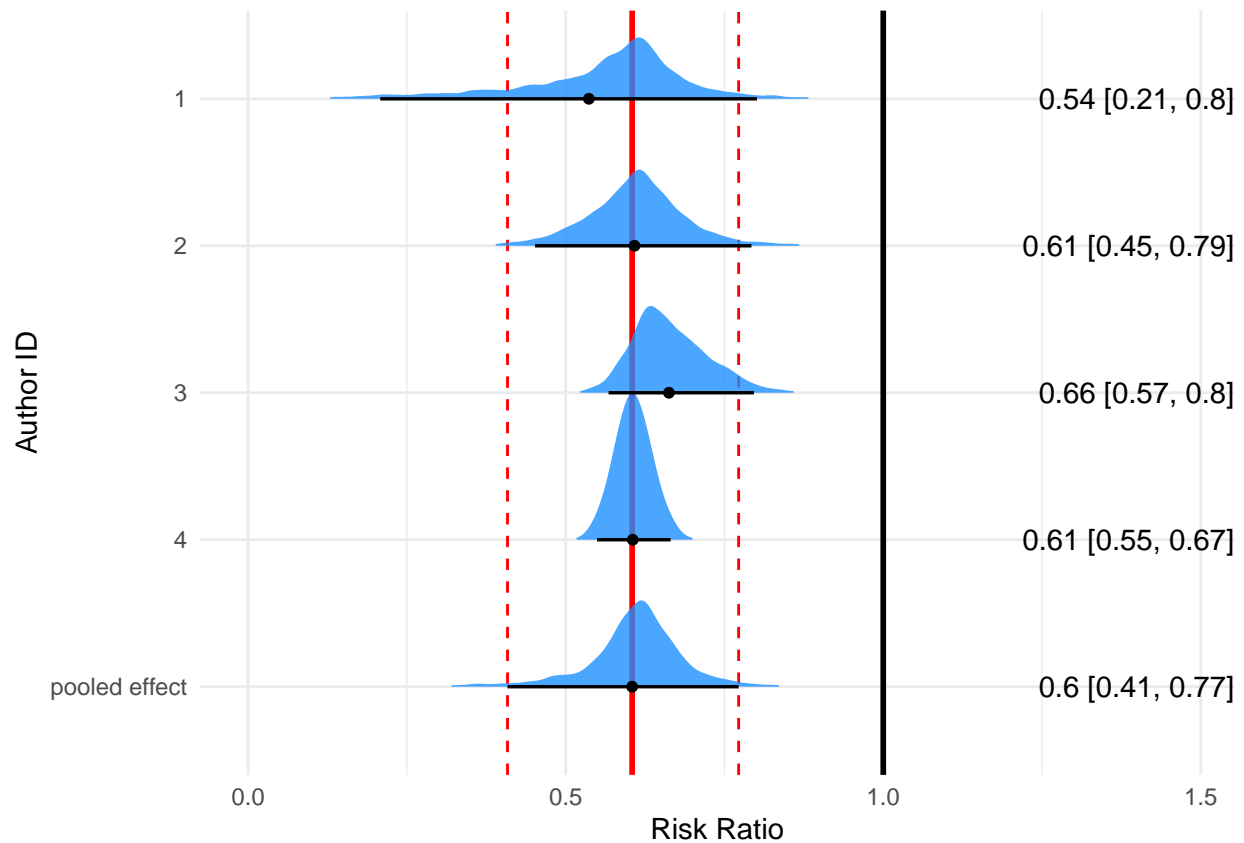
5.Death

```
# Fit Bayesian meta-analysis model
death.bayes <- brm(data = death.es,
                   family = "gaussian",
                   yi | se(sei) ~ 1 + (1 | author_id),
                   prior = priors_half_cauchy,
                   iter = 2000, warmup = 1000, cores = 4, chains = 4,
                   control = list(adapt_delta = 0.99),
                   seed = 18006513)

# Summary of the Bayesian model
summary(death.bayes)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = identity
## Formula: yi | se(sei) ~ 1 + (1 | author_id)
##    Data: death.es (Number of observations: 3)
##   Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup draws = 4000
##
## Multilevel Hyperparameters:
## ~author_id (Number of levels: 3)
```

```
##                   Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)         0.35      0.38     0.01     1.34 1.00      986     1170
##
## Regression Coefficients:
##              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept       -0.22      0.31    -0.82     0.46 1.00     1241     1032
##
## Further Distributional Parameters:
##         Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma       0.00      0.00     0.00     0.00   NA       NA       NA
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```r
# Rhat value for both parameters is 1, signifying convergence
pp_check(death.bayes, ndraws = 100)
```
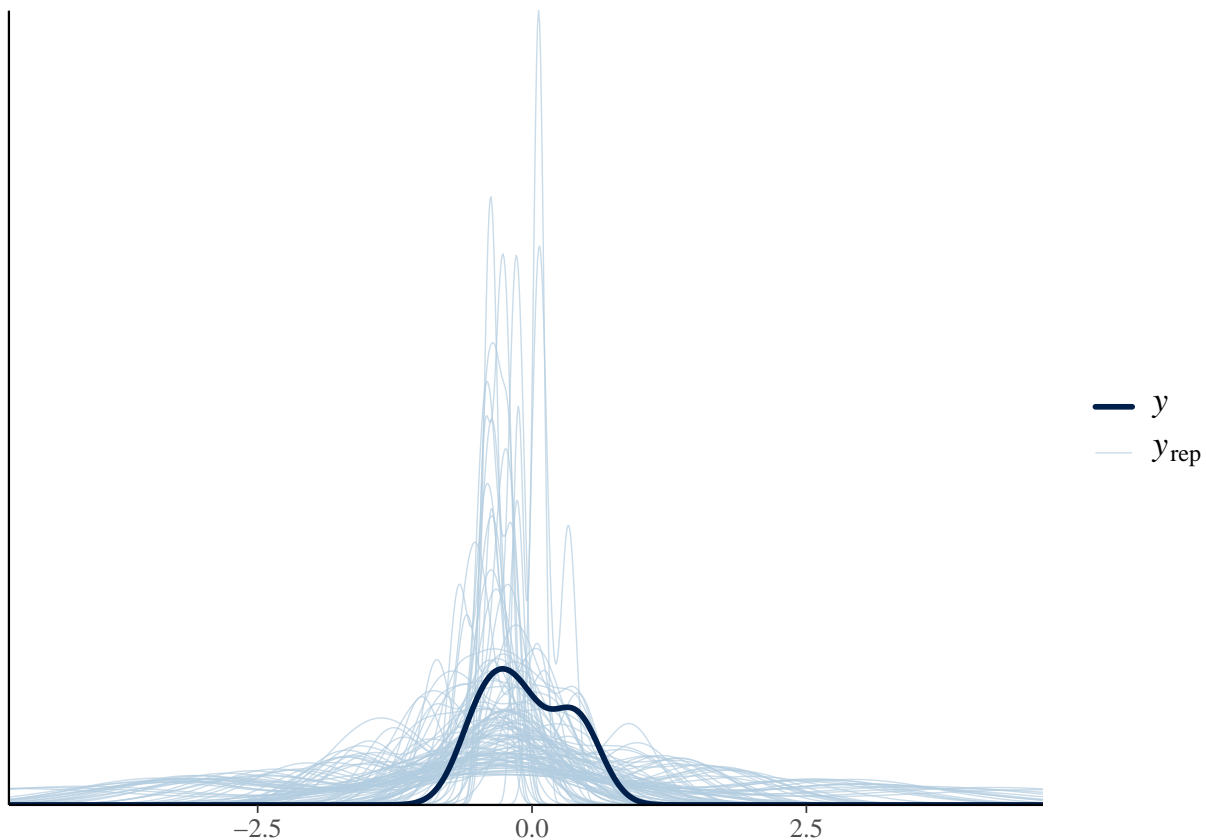


```r
# Prepare Data for Forest Plot
# Spread draws for each study and compute b_Intercept
death.draws <- spread_draws(death.bayes, r_author_id[author_id,], b_Intercept) %>%
  mutate(b_Intercept = r_author_id + b_Intercept) %>%
  mutate(author_id = as.character(author_id))

# Pooled effect
```

```r
death.pooled.draws <- spread_draws(death.bayes, b_Intercept) %>%
  mutate(author_id = "pooled effect")

# Combine individual study and pooled effect data frames
author_levels5 <- c(as.character(1:3), "pooled effect")
death.forest <- bind_rows(death.draws, death.pooled.draws) %>%
  ungroup() %>%
  mutate(author_id = factor(author_id, levels = author_levels5))

# Calculate summary statistics
death.summ <- group_by(death.forest, author_id) %>%
  mean_qi(b_Intercept)

# Exponentiate to get risk ratios
death.forest <- death.forest %>%
  mutate(b_Intercept = exp(b_Intercept))

death.summ <- death.summ %>%
  mutate(b_Intercept = exp(b_Intercept),
         .lower = exp(.lower),
         .upper = exp(.upper))

# Forest plot
ggplot(aes(x = b_Intercept, y = author_id), data = death.forest) +

  # Add vertical lines for pooled effect and CI, and RR = 1
  geom_vline(xintercept = exp(fixef(death.bayes)[1, 1]),
             color = "red", size = 1) +
  geom_vline(xintercept = exp(fixef(death.bayes)[1, 3:4]),
             color = "red", linetype = 2) +
  geom_vline(xintercept = 1, color = "black",
             size = 1) +

  # Add densities
  geom_density_ridges(fill = "dodgerblue",
                      rel_min_height = 0.01,
                      col = NA, scale = 1,
                      alpha = 0.8) +
  geom_pointintervalh(data = death.summ,
                      size = 1) +

  # Add text and labels
  geom_text(data = mutate_if(death.summ, is.numeric, round, 2),
            aes(label = glue("{b_Intercept} [{.lower}, {.upper}]")), x = Inf), hjust = "inward") +
  labs(x = "Risk Ratio", y = "Author ID") +
  theme_minimal() +
  scale_y_discrete(limits = rev(levels(death.summ$author_id))) +
  scale_x_continuous(limits = c(0, 3), breaks = seq(0.5, 2, by = 0.5))
```