Bayesian Model Averaging Meta Analysis

Sophie Berkhout

3/2/2020

```r
library(metaBMA)
library(ggplot2)
library(knitr)
library(kableExtra)
library(papaja)
library(lemon)

ggplot2::theme_set(theme_apa(base_size = 10))
```

```r
default.prior     <- prior(family = "t",
                           param = c(0, 0.707, 1)
                           )

oosterwijk.prior  <- prior(family = "t",
                           param = c(0.35, 0.102, 3)
                           )

vohs.prior        <- prior(family = "norm",
                           param = c(0.3, 0.15)
                           )

priors <- list(default = default.prior,
               oosterwijk = oosterwijk.prior,
               vohs = vohs.prior
               )

# lower truncate priors at zero?
```

```r
.bmaCalculateBFHeterogeneity <- function(prior_models, posterior_models){
# Returns the heterogeneity Bayes factor

  postOdds <- (posterior_models["random_H0"] + posterior_models["random_H1"]) /
              (posterior_models["fixed_H0"] + posterior_models["fixed_H1"])
  priorOdds <- (prior_models[3] + prior_models[4]) / (prior_models[1] + prior_models[2])
  BFheterogeneity <- postOdds/priorOdds
  return(BFheterogeneity)
}

# BFrf or BFfr?
```

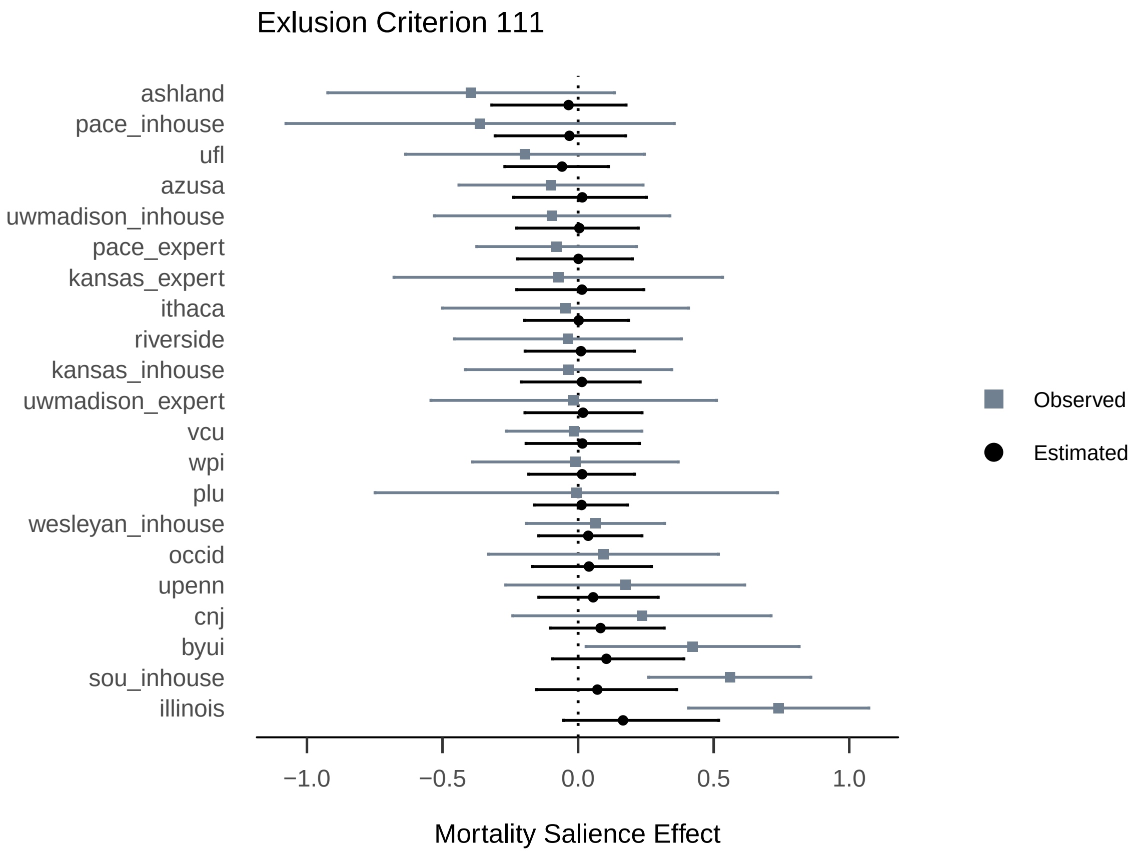
```r
.runRema <- function(y, SE, priors, extrainfo = FALSE){
# Returns a list containing:
# BFmu: dataframe with 1 row/3columns (BF per prior)
# BFtau: dataframe with 1 row/3columns (BF per prior)
# est: list containing 3 dataframes per prior
#      with effect size and 95% credible interval estimates
```

```
  BFmu <- BFtau <- data.frame(default = 0,
                              oosterwijk = 0,
                              vohs = 0
                              )

  empty.dataframe <- data.frame(est.es = numeric(),
                                est.lower = numeric(),
                                est.upper = numeric()
                                )

  est <- list(default = empty.dataframe,
              oosterwijk = empty.dataframe,
              vohs = empty.dataframe
              )

  for(i in 1:length(priors)){
    rema <- meta_bma(y, SE, d = priors[[i]])
    BFmu[, i] <- 1/rema$inclusion$incl.BF # BF in favour of the null model
    if(extrainfo){
      BFtau[i] <- .bmaCalculateBFHeterogeneity(rema$prior_models, rema$posterior_models)
      est.es <- rstan::summary(rema$meta$random$stanfit_dstudy)$summary[3:(length(y) + 2), "mean"]
      est.lower <- rstan::summary(rema$meta$random$stanfit_dstudy)$summary[3:(length(y) + 2), "2.5%"]
      est.upper <- rstan::summary(rema$meta$random$stanfit_dstudy)$summary[3:(length(y) + 2), "97.5%"]
      est[[i]] <- rbind(est[[i]], cbind(est.es, est.lower, est.upper))
    }
  }
  if(!extrainfo){
    return(BFmu)
  } else {
    remaResult <- list(BFmu = BFmu,
                       BFtau = BFtau,
                       est = est
                       )
    return(remaResult)
  }
}
```

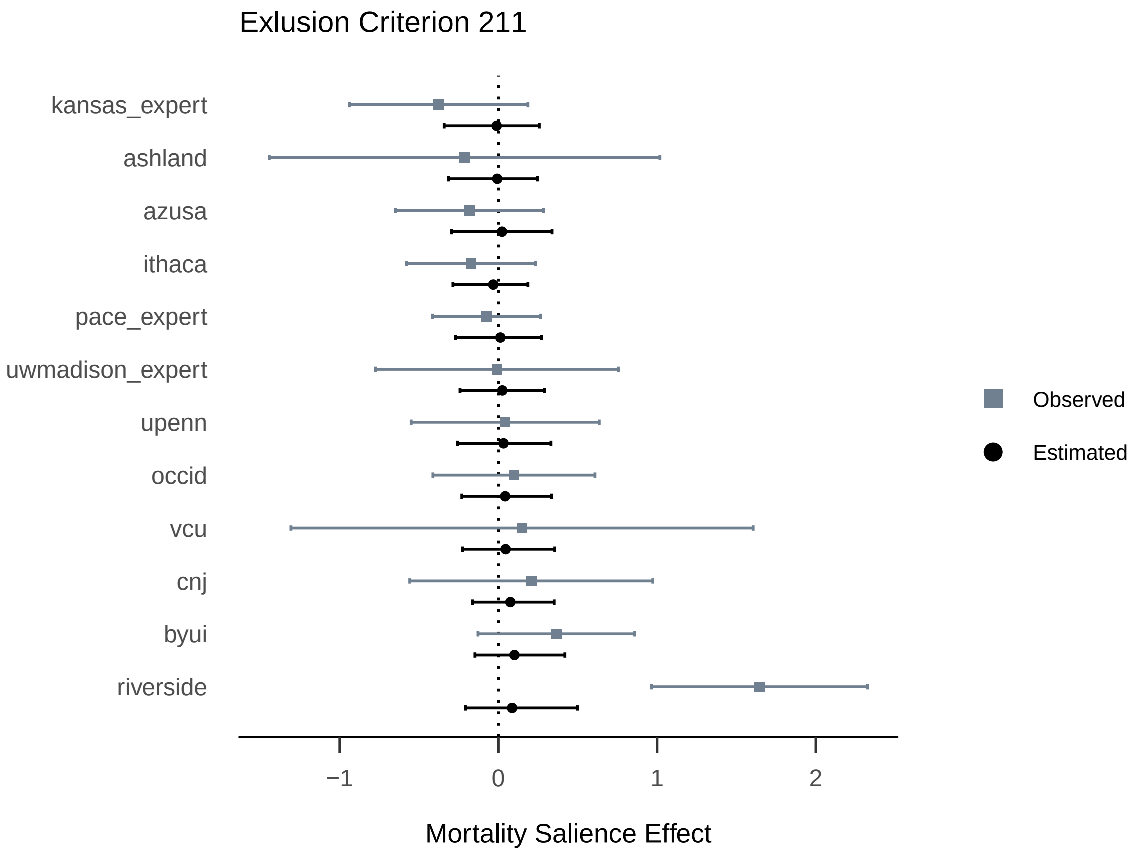
```
remaForest <- function(obs.data, est.data){
# Returns a forest plot with observed and estimated effects

I <- nrow(obs.data)
ord <- order(obs.data$yi)
es <- obs.data$yi[ord]
est.es <- est.data[ord, "est.es"]
est.lower <- est.data[ord, "est.lower"]
est.upper <- est.data[ord, "est.upper"]
studyLabels <- obs.data[ord, "source"]

lower <- es - qnorm(1.95/2) * obs.data$sei
upper <- es + qnorm(1.95/2) * obs.data$sei

y.obs <- I:1
y.est <- rev(seq(.6, I - .4, 1))
dfBoth <- data.frame(es = c(es, est.es),
                     y = c(y.obs, y.est),
                     studyLabels = c(studyLabels, studyLabels),
                     lower = c(lower, est.lower), upper = c(upper, est.upper),
                     g = rep(c("Observed", "Estimated"), each = I)
                     )


plot <- ggplot(dfBoth, aes(x = es, y = y)) +
        geom_vline(xintercept = 0, linetype = "dotted") +
        geom_point(aes(shape = as.factor(dfBoth$g), colour = as.factor(dfBoth$g))) +
        geom_errorbarh(aes(xmin = dfBoth$lower, xmax = dfBoth$upper, colour = as.factor(dfBoth$g)),
                        height = .1, show.legend = FALSE) +
        scale_y_continuous(breaks = I:1, labels = as.character(studyLabels),
                                    expand = c(0, 0.5)) +
        scale_color_manual("", values = c("black", "slategrey"),
                                    labels = c(gettext("Estimated"), gettext("Observed"))) +
        scale_shape_manual("", values = c(16, 15)) +
        guides(shape = guide_legend(reverse=TRUE, override.aes = list(size=3)),
               colour = guide_legend(reverse=TRUE))  +
        xlab("Mortality Salience Effect") +
        ylab(" ") +
        theme(axis.line.y = element_blank(),
              axis.ticks.y = element_blank())

return(plot)
}
```

Original Analysis

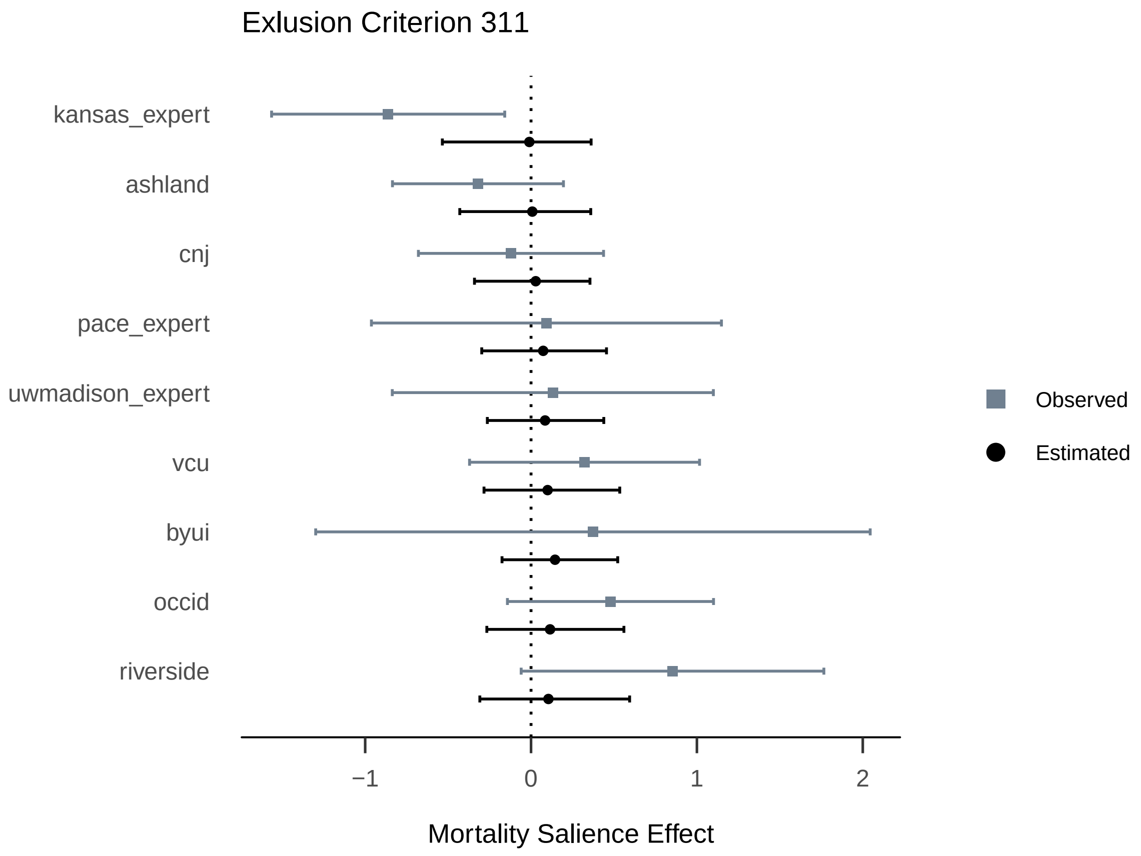
```
m.111 <- read.csv2("data/metaset_111.csv", header = T)
m.211 <- read.csv2("data/metaset_211.csv", header = T)
m.311 <- read.csv2("data/metaset_311.csv", header = T)

rema.111 <- .runRema(m.111$yi, m.111$sei, priors, TRUE)
rema.211 <- .runRema(m.211$yi, m.211$sei, priors, TRUE)
rema.311 <- .runRema(m.311$yi, m.311$sei, priors, TRUE)

remaForest(m.111, rema.111$est$default) + ggtitle("Exlusion Criterion 111")
```



```
remaForest(m.211, rema.211$est$default) + ggtitle("Exlusion Criterion 211")
```



```
remaForest(m.311, rema.311$est$default) + ggtitle("Exlusion Criterion 311")
```

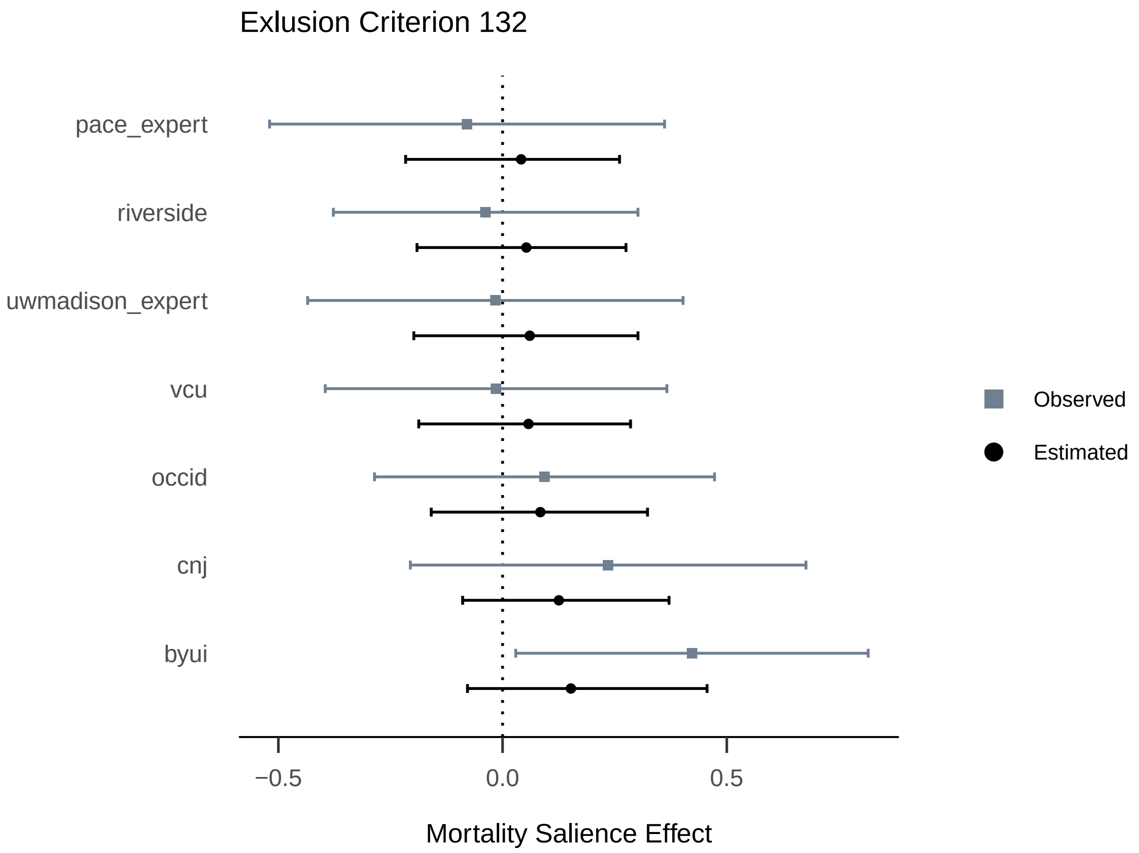


```
BFs.og <- cbind(rbind(rema.111$BFmu, rema.211$BFmu, rema.311$BFmu),
                rbind(rema.111$BFtau, rema.211$BFtau, rema.311$BFtau))
BFs.og <- round(BFs.og, 2)
rownames(BFs.og) <- paste("Exclusion Criterion", c(111, 211, 311))
subhead <- rep(c("Default", "Oosterwijk", "Vohs"), 2)
colnames(BFs.og) <- subhead
colhead <- list("Effect Size BF" = c(2, 4),
                "Heterogeneity BF" = c(5, 7))

apa_table(BFs.og, escape = FALSE, caption = "Model-averaged Bayes factors for key analyses.",
          note = "All Bayes factors are reported in favor of the null model.",
          col_spanners = colhead)
```

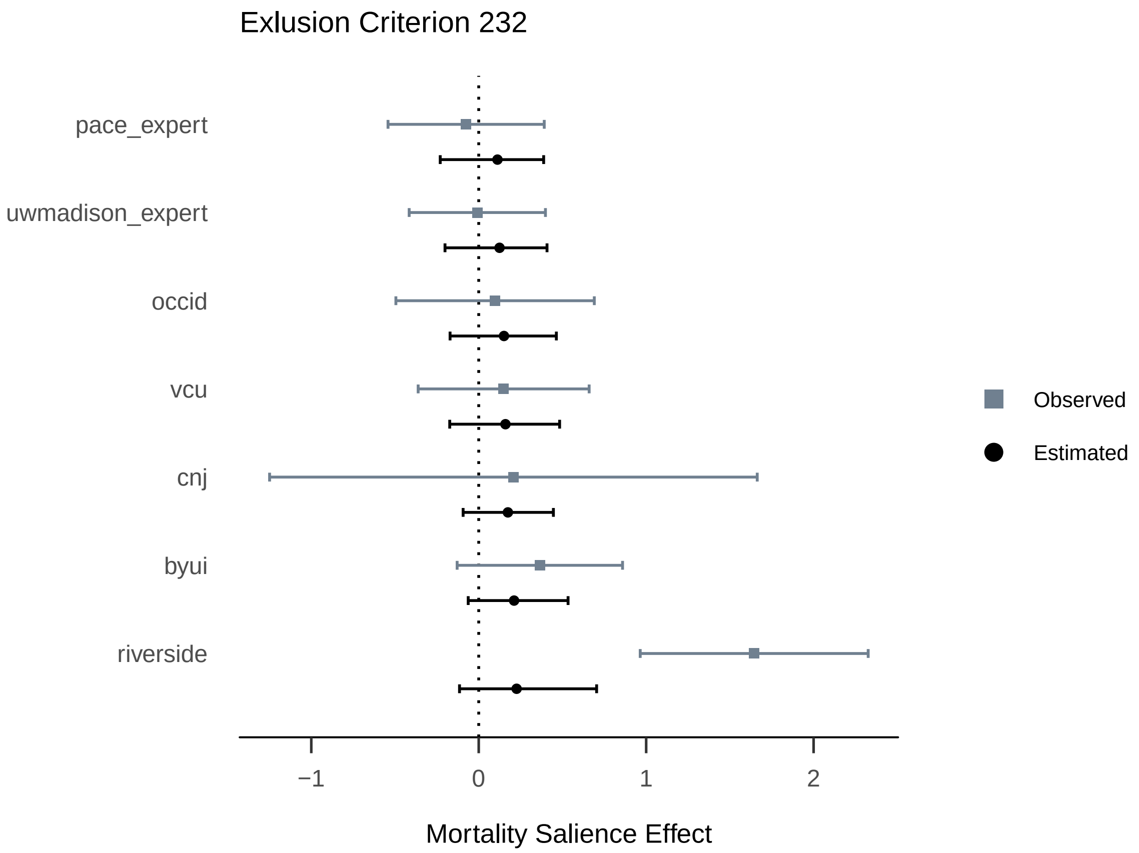
```
##
##
## \begin{table}[tbp]
## \begin{center}
## \begin{threeparttable}
## \caption{\label{tab:unnamed-chunk-7}Model-averaged Bayes factors for key analyses.}
## \begin{tabular}{lllllll}
## \toprule
##  & \multicolumn{3}{c}{Effect Size BF} & \multicolumn{3}{c}{Heterogeneity BF} \\
## \cmidrule(r){2-4} \cmidrule(r){5-7}
##  & \multicolumn{1}{c}{Default} & \multicolumn{1}{c}{Oosterwijk} & \multicolumn{1}{c}{Vohs} & \multic
## \midrule
## Exclusion Criterion 111 & 17.40 & 37.66 & 14.57 & 0.44 & 0.43 & 0.44\\
## Exclusion Criterion 211 & 10.10 & 15.14 & 6.63 & 0.41 & 0.42 & 0.42\\
## Exclusion Criterion 311 & 6.01 & 3.94 & 2.45 & 0.56 & 0.57 & 0.57\\
## \bottomrule
## \addlinespace
## \end{tabular}
## \begin{tablenotes}[para]
## \normalsize{\textit{Note.} All Bayes factors are reported in favor of the null model.}
## \end{tablenotes}
## \end{threeparttable}
## \end{center}
## \end{table}
```

Analysis for the Main Claim of the Comment

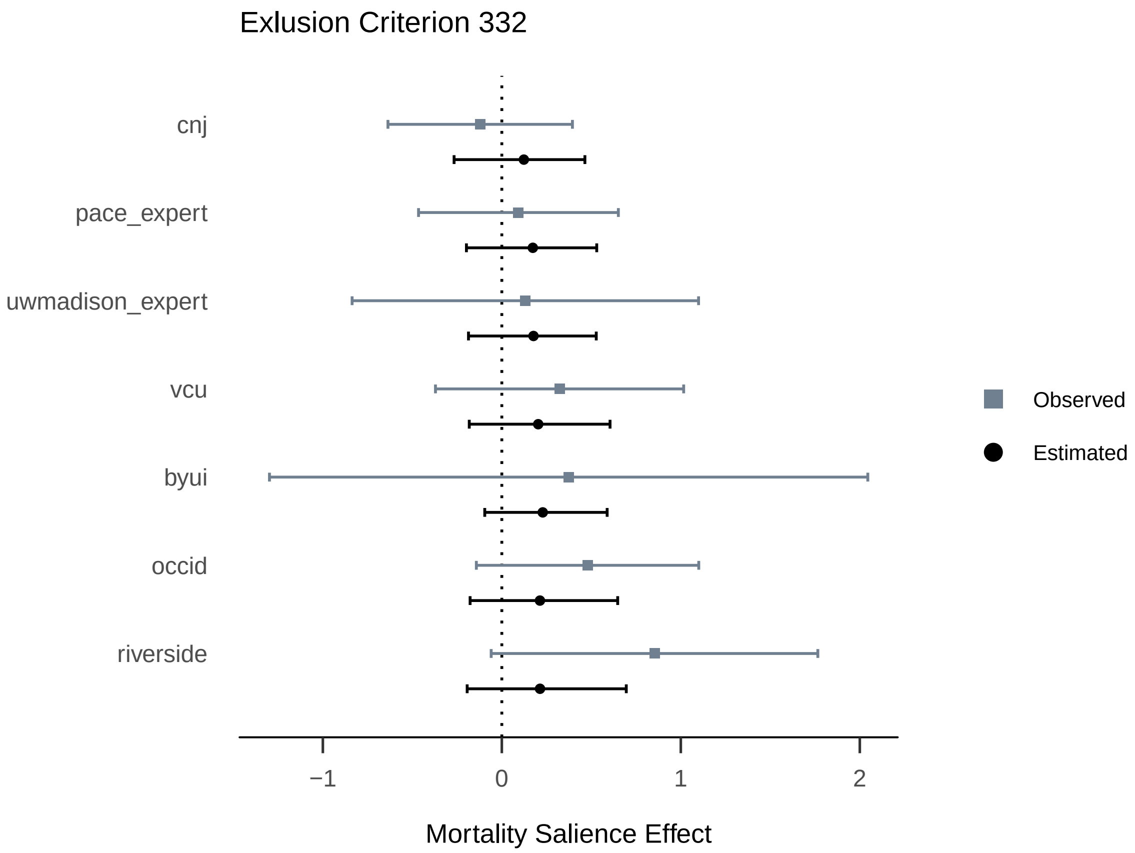
```
m.132 <- read.csv2("data/metaset_132.csv", header = T)
m.232 <- read.csv2("data/metaset_232.csv", header = T)
m.332 <- read.csv2("data/metaset_332.csv", header = T)

rema.132 <- .runRema(m.132$yi, m.132$sei, priors, TRUE)
rema.232 <- .runRema(m.232$yi, m.232$sei, priors, TRUE)
rema.332 <- .runRema(m.332$yi, m.332$sei, priors, TRUE)

remaForest(m.132, rema.132$est$default) + ggtitle("Exlusion Criterion 132")
```



```
remaForest(m.232, rema.232$est$default) + ggtitle("Exlusion Criterion 232")
```



```
remaForest(m.332, rema.332$est$default) + ggtitle("Exlusion Criterion 332")
```



```
BFs.cmnt <- cbind(rbind(rema.132$BFmu, rema.232$BFmu, rema.332$BFmu),
                  rbind(rema.132$BFtau, rema.232$BFtau, rema.332$BFtau))
BFs.cmnt <- round(BFs.cmnt, 2)
rownames(BFs.cmnt) <- paste("Exclusion Criterion", c(132, 232, 332))
colnames(BFs.cmnt) <- subhead

apa_table(BFs.cmnt, escape = FALSE, caption = "Model-averaged Bayes factors for comment analyses.",
          col_spanners = colhead)
```

```
##
##
## \begin{table}[tbp]
## \begin{center}
## \begin{threeparttable}
## \caption{\label{tab:unnamed-chunk-9}Model-averaged Bayes factors for comment analyses.}
## \begin{tabular}{lllllll}
## \toprule
##  & \multicolumn{3}{c}{Effect Size BF} & \multicolumn{3}{c}{Heterogeneity BF} \\
## \cmidrule(r){2-4} \cmidrule(r){5-7}
##  & \multicolumn{1}{c}{Default} & \multicolumn{1}{c}{Oosterwijk} & \multicolumn{1}{c}{Vohs} & \multic
## \midrule
## Exclusion Criterion 132 & 6.50 & 5.78 & 2.73 & 0.40 & 0.40 & 0.40\\
## Exclusion Criterion 232 & 2.65 & 0.92 & 0.67 & 0.49 & 0.50 & 0.48\\
## Exclusion Criterion 332 & 2.63 & 0.74 & 0.63 & 0.54 & 0.54 & 0.52\\
## \bottomrule
## \end{tabular}
## \end{threeparttable}
## \end{center}
## \end{table}
```

All Other Variants

```
sets <- c(112, 212, 312, 121, 221, 321, 122, 222, 322, 131, 231, 331)
rema.other <- data.frame(metaset = character(),
                         default = numeric(),
                         oosterwijk = numeric(),
                         vohs = numeric()
                         )
for(i in sets){
  path.m <- paste0("data/metaset_", i, ".csv")
  m <- read.csv2(path.m, header = T)
  rema <- .runRema(m$yi, m$sei, priors, FALSE)
  rema.other <- rbind(rema.other, cbind(rema))
}
```

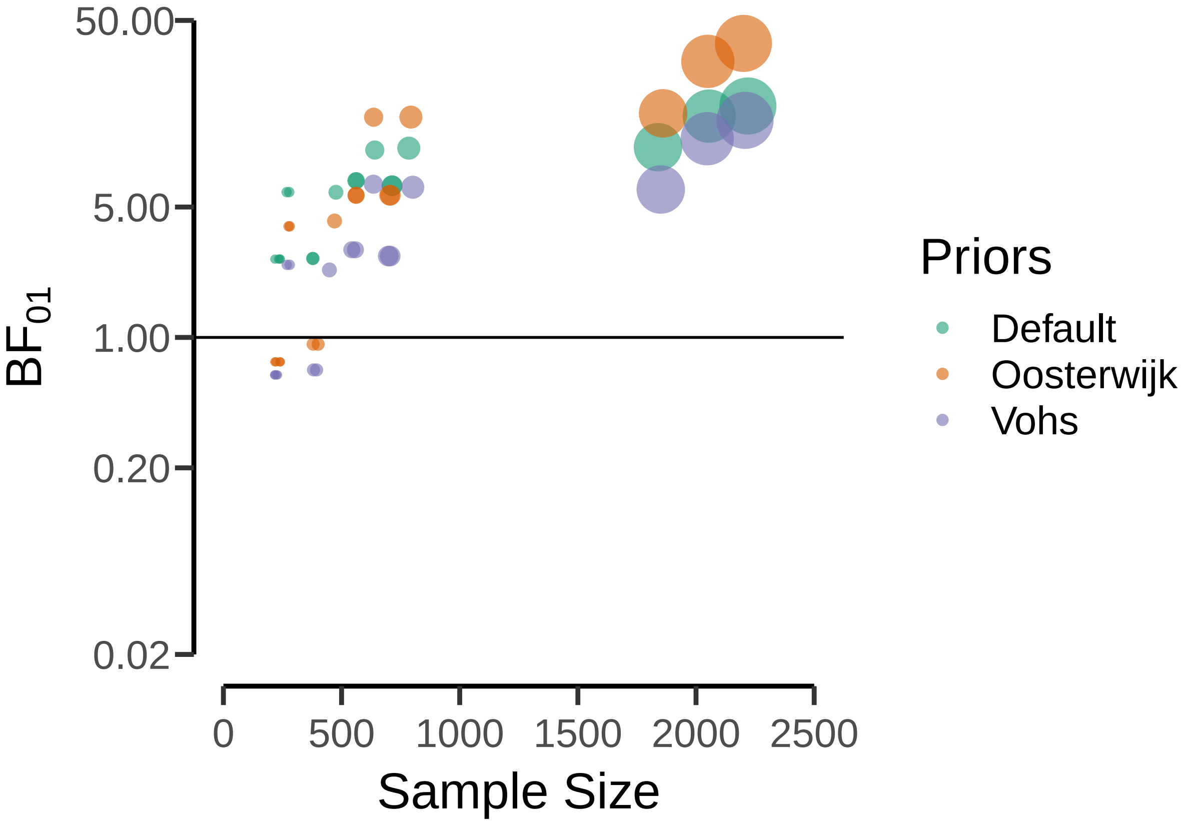
```
rema.other <- round(rema.other, 2)
rownames(rema.other) <- paste("Exclusion Criterion", sets)
colnames(rema.other) <- c("Default", "Oosterwijk", "Vohs")

apa_table(rema.other, escape = FALSE, caption = "Model-averaged Bayes factors for all other analyses.")
```

```
##
##
## \begin{table}[tbp]
## \begin{center}
## \begin{threeparttable}
## \caption{\label{tab:unnamed-chunk-11}Model-averaged Bayes factors for all other analyses.}
## \begin{tabular}{llll}
## \toprule
##  & \multicolumn{1}{c}{Default} & \multicolumn{1}{c}{Oosterwijk} & \multicolumn{1}{c}{Vohs}\\
## \midrule
## Exclusion Criterion 112 & 10.34 & 15.16 & 6.38\\
## Exclusion Criterion 212 & 6.00 & 4.21 & 2.30\\
## Exclusion Criterion 312 & 6.01 & 3.94 & 2.45\\
## Exclusion Criterion 121 & 15.34 & 30.19 & 11.62\\
## Exclusion Criterion 221 & 6.92 & 5.78 & 2.95\\
## Exclusion Criterion 321 & 2.63 & 0.74 & 0.63\\
## Exclusion Criterion 122 & 6.50 & 5.78 & 2.73\\
## Exclusion Criterion 222 & 2.65 & 0.92 & 0.67\\
## Exclusion Criterion 322 & 2.63 & 0.74 & 0.63\\
## Exclusion Criterion 131 & 10.46 & 15.88 & 6.21\\
## Exclusion Criterion 231 & 6.92 & 5.78 & 2.95\\
## Exclusion Criterion 331 & 2.63 & 0.74 & 0.63\\
## \bottomrule
## \end{tabular}
## \end{threeparttable}
## \end{center}
## \end{table}
```

Combining Evidence

```r
sets <- c(111, 211, 311, 132, 232, 332, 112, 212, 312, 121, 221, 321, 122, 222, 322, 131, 231, 331)
ssize <- data.frame("metaset" = character(),
                         "n.study" = numeric()
                         )
for(i in sets){
  path.m <- paste0("data/reanalysis_", i, ".csv")
  m <- read.csv2(path.m, header = T)
  N <- nrow(m)
  ssize <- rbind(ssize, c(i, N))
}
```

```r
rema.all <- rbind(BFs.og[, 1:3]
                  , BFs.cmnt[, 1:3]
                  , rema.other
                  )
rema.all <- cbind(rema.all, "N" = ssize[, 2])
rema.all <- cbind(rema.all, "Crit" = rownames(rema.all))
result <- tidyr::gather(rema.all, prior, bf, Default:Vohs, factor_key = TRUE)
result$Nscaled <- result$N/mean(result$N) * 3

xrange <- c(0, 2500)
yrange <- c(1/50, 50) # log scale

# Plot with n participants included on the x-axis, BF01 on the y-axis, colors for
# participant exclusion sets and bubble sizes for number of studies included
ggplot(aes(x = N, y = bf, size = Nscaled, color = as.factor(prior)), data = result) +
  geom_jitter(alpha = 0.6, width = 15) +
  scale_y_continuous(trans='log', limits = yrange, breaks = c(1/50, 1/5, 1, 5, 50)) +
  scale_size_identity() +
  scale_colour_brewer(palette = 'Dark2') +
  labs(x = 'Sample Size', y = expression(BF["01"]), color = "Priors") +
  xlim(xrange) +
  geom_hline(yintercept = 1) +
  theme_apa() +
  theme_classic(base_size = 19) +
  theme(axis.line=element_line()) +
  coord_capped_cart(gap = 0.1, bottom = 'none', left = 'none') + #caps the axes so they don't touch
  theme(axis.title.x = element_text(vjust=-0.5),
        axis.text.x = element_text(vjust=-0.5),
        axis.text.y = element_text(hjust=.8,),
        axis.ticks.length = unit(0.25, 'cm'),

        )
```


50.00
5.00
1.00
0.20
0.02
BF_{01}
0
500
1000
1500
2000
2500
Sample Size
Priors
Default
Oosterwijk
Vohs