Literature Coding Analysis for 'Characterising the use of internal meta-analyses and assessing their impact'

Mandy Norrbo & Lisa DeBruine

27/02/2020

Contents

Dependencies

```
library(tidyverse)
```

Session info

```
sessionInfo()
```

```
## R version 3.6.1 (2019-07-05)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 17763)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United Kingdom.1252
## [2] LC_CTYPE=English_United Kingdom.1252
## [3] LC_MONETARY=English_United Kingdom.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United Kingdom.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] forcats_0.4.0   stringr_1.4.0   dplyr_0.8.3     purrr_0.3.3
## [5] readr_1.3.1     tidyr_1.0.2     tibble_2.1.3    ggplot2_3.2.1
## [9] tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
##  [1] tidyselect_0.2.5 xfun_0.9         haven_2.2.0      lattice_0.20-38
##  [5] colorspace_1.4-1 vctrs_0.2.2      generics_0.0.2   htmltools_0.4.0
##  [9] yaml_2.2.0       rlang_0.4.3      pillar_1.4.3     withr_2.1.2
## [13] glue_1.3.1       DBI_1.0.0        dbplyr_1.4.2     modelr_0.1.5
## [17] readxl_1.3.1     lifecycle_0.1.0  munsell_0.5.0    gtable_0.3.0
## [21] cellranger_1.1.0 rvest_0.3.5      evaluate_0.14    knitr_1.24
## [25] fansi_0.4.1      broom_0.5.2      Rcpp_1.0.3       scales_1.1.0
## [29] backports_1.1.5  jsonlite_1.6     fs_1.3.1         hms_0.5.3
## [33] digest_0.6.23    stringi_1.4.4    grid_3.6.1       cli_2.0.1
## [37] tools_3.6.1      magrittr_1.5     lazyeval_0.2.2   crayon_1.3.4
## [41] pkgconfig_2.0.3  xml2_1.2.2       reprex_0.3.0     lubridate_1.7.4
## [45] assertthat_0.2.1 rmarkdown_1.15   httr_1.4.1       rstudioapi_0.10
## [49] R6_2.4.1         nlme_3.1-141     compiler_3.6.1
```

Import data

```
dat <- read_csv("quantdata.csv")
```

Analysis

Data wrangling

Creating a dataframe that contains only data for comparing internal meta-analysis articles and matched control articles.

```
tidydat <- dat %>%
  #selecting columns of interest in analysis
  select(-c("TITLE", "DOI", "EMPIRICAL", "SIGTEST", "COMMENT")) %>%
  #selecting the studies that should be included in the control vs minimeta analysis
  filter(INCLUDE == "yes") %>%
  #filtering out empty columns
  select(-"PVAL8", -"PVAL9", -"PVAL10", -"PVAL11",
         -"PVAL12", -"N8", -"N9", -"N10", -"N11", -"N12") %>%
  gather(key = "PVALNO", value = "PVALUES",
         PVAL1,PVAL2,PVAL3,PVAL4,PVAL5,PVAL6, PVAL7)
```

Descriptives

Prevalence

Calculating how many articles used internal meta-analysis out of all coded empirical articles

```
nmini <- dat %>%
  filter(EMPIRICAL == "yes") %>%
  count(USES_MINIMETA == "yes") %>%
  arrange(n)
```

14 out of 66 empirical research articles published between May and September 2019 in the Journal of Experimental Social Psychology used an internal meta-analysis — that is 21%.

Multi-study papers

Calculating how many articles included more than one study

```
#no of articles with more than 1 study
multidat <- dat %>%
  select(STUDY_NO) %>%
  filter(STUDY_NO > 1) %>%
  nrow()

#no of articles with just 1 study
singledat <- dat %>%
  select(STUDY_NO) %>%
  filter(STUDY_NO == 1) %>%
  nrow()
```

59 out of 66 empirical articles reported the findings of more than one study, only 7 reported the findings from a single study.

Aggregated studies

Calculating how many studies were combined in internal meta-analysis articles

```
#selecting articles that used minimeta
minidat <- tidydat %>%
  filter(USES_MINIMETA == "yes")
```

```r
#number of individual studies that were aggregated
nstudymini <- minidat  %>%
  summarise(mean = mean(META_STUDY_NO),
            median = median(META_STUDY_NO),
            sd = sd(META_STUDY_NO))
```

The number of studies combined in the mini meta-analyses coded in this overview range from 2 to 7, with a mean of 4.2, median of 4, and with a standard deviation of 1.55.

Sample sizes

```r
#sample size dataframe for plotting & descriptives
ndat <- dat %>%
  filter(INCLUDE == "yes", USES_MINIMETA == "yes") %>%
  select(c("N1", "N2", "N3", "N4", "N5", "N6", "N7", "USES_MINIMETA", "STUDY_ID")) %>%
  gather(key = "NSTUDY", value = "NVALUES",
         N1, N2, N3, N4, N5, N6, N7)
```

Sample sizes within single studies ranged from 40 to 988, with a mean size of 325 and median size of 282.

How many p > 0.05?

```r
#how many p-values > 0.05 grouped by minimeta use (yes/no)
nullp <- tidydat %>%
  filter(PVALUES > 0.05) %>%
  group_by(USES_MINIMETA) %>%
  summarise(n = n())

#how many individual studies in total in 20 articles
studytot <- tidydat %>%
  filter(PVALUES != "") %>%
  nrow()
```

Out of the 20 articles and 80 individual studies, 13 p-values were nonsignificant. Out of these, 11 were in the articles that used internal meta-analyses, while 2 were in the matched control articles that did not end up using an internal meta-analysis.

Pre-registrations

```r
#dataframe with prereg information
prereg <- dat %>%
  filter(INCLUDE == "yes", USES_MINIMETA == "yes") %>%
  select(STUDY_ID, STUDY_NO,PRE_REG)

#percentage pre-registered
percreg <- round(sum(prereg$PRE_REG)/sum(prereg$STUDY_NO),2)*100
#total no. pre-registered
totalreg <- sum(prereg$PRE_REG)
#out of how many studies
outof <- sum(prereg$STUDY_NO)
```

Within the internal meta-analysis articles, 8 out of 47, , that is 17%, were reported to have been pre-registered. Only 1 article (82_2) seemed to have pre-registered all studies within the internal meta-analysis. However,

the internal meta-analysis itself did not seem to be pre-registered and thus, when to run it and which studies to include was still flexible.

How many do not report p-values?

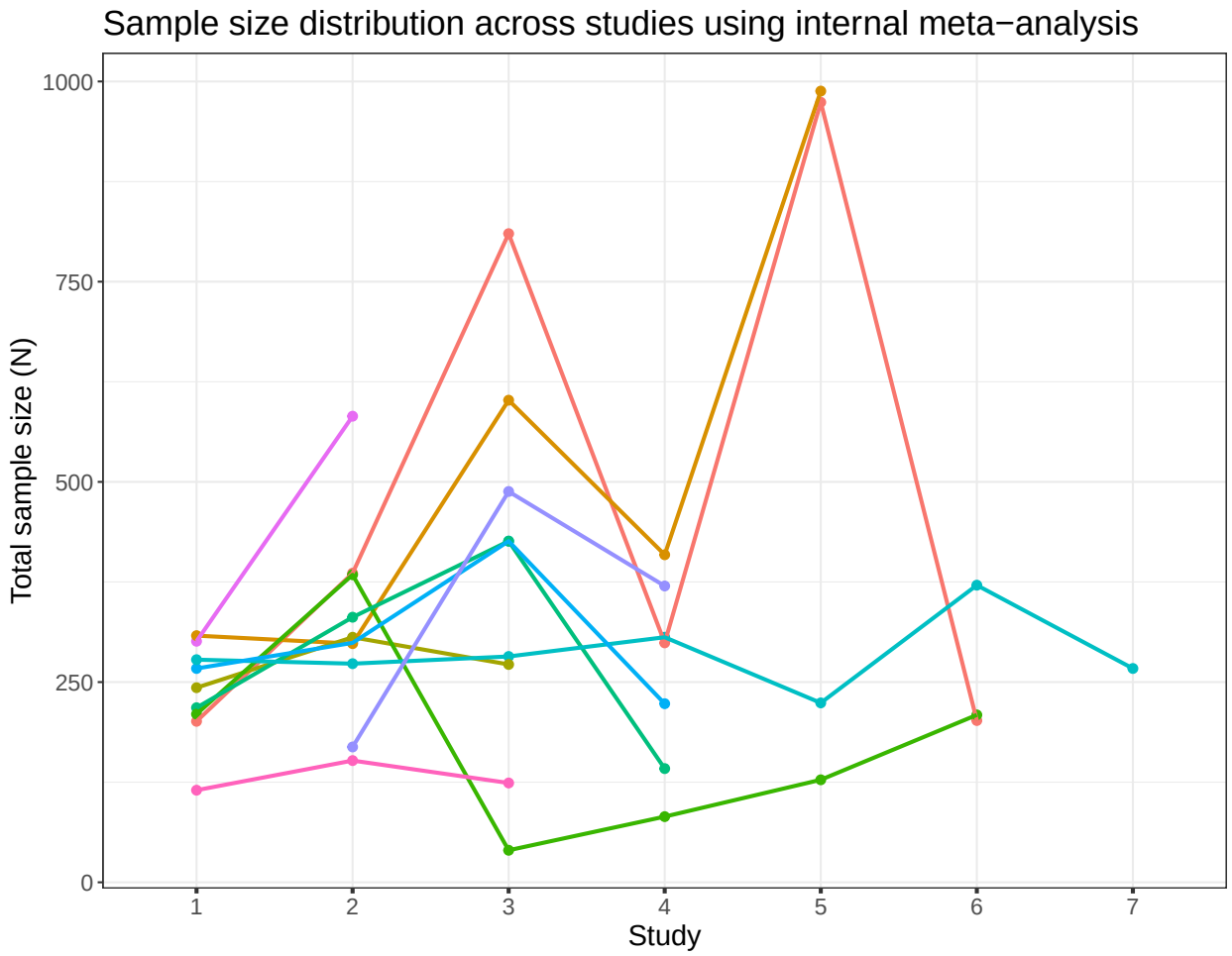
```
#articles without p-values were coded as "not sure" in "IS_MINI_SIG" variable
nopval <- dat %>%
  filter(IS_MINI_SIG == "not sure") %>%
  nrow()
```

Out of the 14 coded articles using an internal meta-analysis, 3 did not report an overall significance level (i.e. p-value) for the meta-analysis, instead reporting only effect sizes and confidence intervals.

Visualisation

sample sizes

```
g2 <- ggplot(ndat, aes(x=NSTUDY, y=NVALUES)) +
  geom_point(aes(colour = STUDY_ID), size = 2) +
  geom_line(aes(group = STUDY_ID, colour = STUDY_ID), size = 1) +
  theme(axis.text.x  = element_text(size = 5)) +
  theme_bw() +
  theme(legend.position = "none") +
  scale_x_discrete(labels=c("N1" = "1", "N2" = "2",
                            "N3" = "3", "N4" = "4",
                            "N5"= "5", "N6" = "6", "N7" = "7"), name = "Study") +
  scale_y_continuous(name = "Total sample size (N)") +
  labs(title = "Sample size distribution across studies using internal meta-analysis") +
  theme(axis.title = element_text(size = 15),
        title = element_text(size = 15),
        axis.text = element_text(size = 10),
        axis.ticks.x = element_line(size = 1),
        axis.text.x =element_text(size=12),
        axis.text.y = element_text(size=12),
        legend.text = element_text(size = 12))
g2
```




```
#ggsave("minimeta_n_plot.png", width = 10, height = 6)
```

p-values

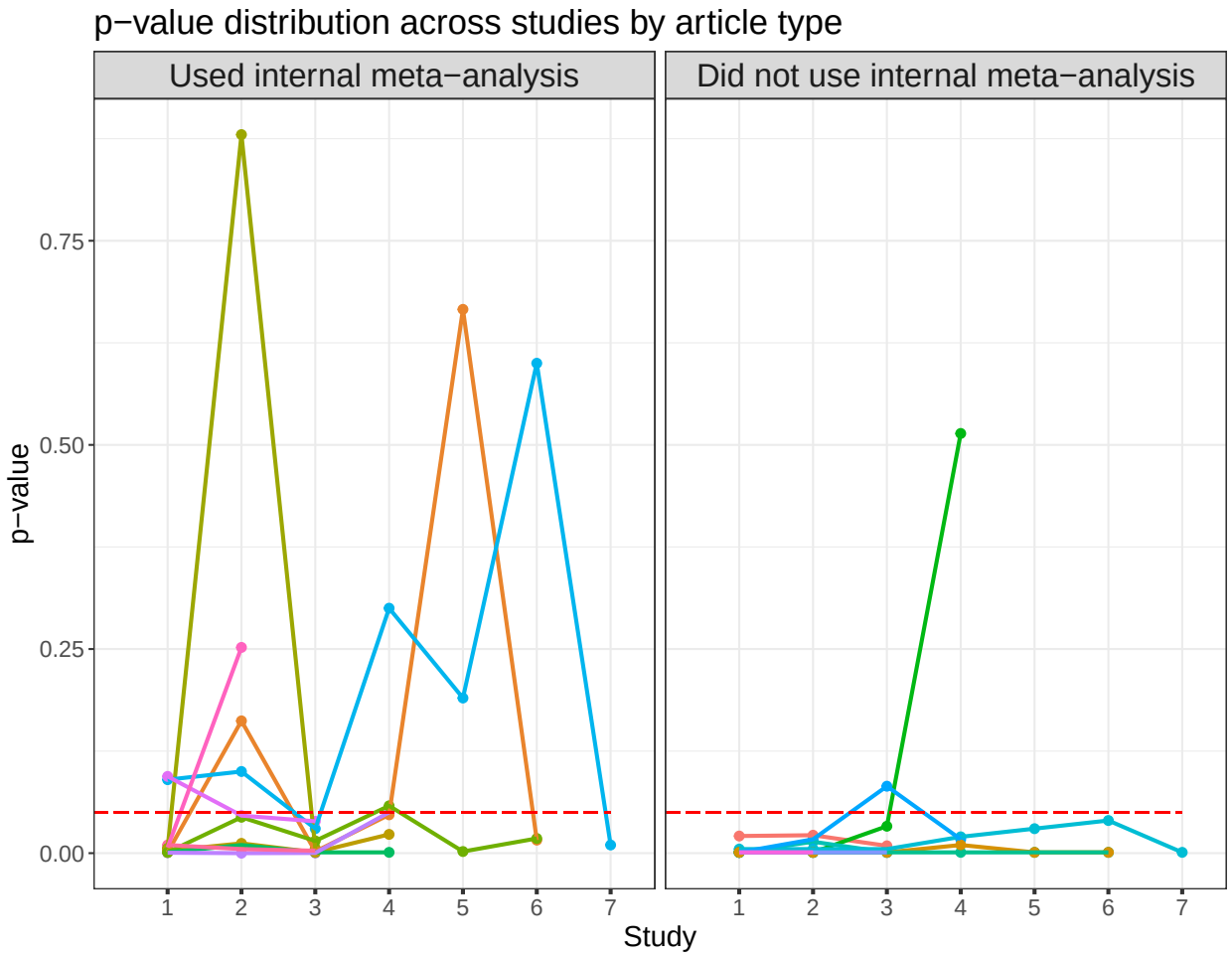
A figure comparing internal meta-analyses to their matched controls regarding the distribution of p-values.

```
#labels for plot
supp.labs <- c('no' = "Used internal meta-analysis",
               'yes' = "Did not use internal meta-analysis")

g1 <- ggplot(tidydat, aes(x=PVALNO, y=PVALUES)) +
  geom_point(aes(colour = STUDY_ID), size = 2) +
  geom_line(aes(group = STUDY_ID, colour = STUDY_ID), size = 1) +
  geom_segment(aes(x = 0, xend = 7, y =.05, yend = 0.05),
               colour = "red", size = 0.5, linetype = "longdash")  +
  theme(axis.text.x  = element_text(size = 5)) +
  theme_bw() +
  theme(legend.position = "none") +
  scale_x_discrete(labels=c("PVAL1" = "1", "PVAL2" = "2",
                            "PVAL3" = "3", "PVAL4" = "4",
                            "PVAL5"= "5", "PVAL6" = "6",
                            "PVAL7" = "7"),
                   name = "Study") +
```

```
  scale_y_continuous(name = "p-value") +
  labs(title = "p-value distribution across studies by article type") +
  theme(axis.title = element_text(size = 15),
        title = element_text(size = 15),
        axis.text = element_text(size = 10),
        axis.ticks.x = element_line(size = 1),
        axis.text.x =element_text(size=12),
        axis.text.y = element_text(size=12),
        legend.text = element_text(size = 12),
        strip.text.x = element_text(size = 17)) +
  facet_wrap(~CONTROL, labeller = as_labeller(supp.labs))

g1
```




```
#ggsave("p_dist_plot.png", width = 10, height = 7)
```