Simpson forestplot

Clémence Métayer metayer.clemence@yahoo.com (main contributor)
Marta Avalos Fernandez marta.avalos-fernandez@u-bordeaux.fr

2022-07-19

Data import

```
library(metafor)
library(forestplot)
library(lme4)
library(lmerTest)

data_study <- read.csv("data_forest_plot_difference.csv", sep = ";",
    dec = ".")
rownames(data_study) <- c("Den16", "Li17", "Hua20", "Kim17",
    "Pra12", "Fil12", "Sor20", "Buy17")
```

Computation of differences and standard deviations with escalc

```
diff_healthy_stable <- escalc(measure = "SMDH", m1i = Mean_healthy,
    sd1i = SD_healthy, n1i = Size_healthy, m2i = Mean_stable,
    sd2i = SD_stable, n2i = Size_stable, data = data_study)
Sum_HS <- summary(diff_healthy_stable)

diff_healthy_exacerbated <- escalc(measure = "SMDH", m1i = Mean_healthy,
    sd1i = SD_healthy, n1i = Size_healthy, m2i = Mean_exacerbated,
    sd2i = SD_exacerbated, n2i = Size_exacerbated, data = data_study)
Sum_HE <- summary(diff_healthy_exacerbated)

diff_healthy_diseased <- escalc(measure = "SMDH", m1i = Mean_healthy,
    sd1i = SD_healthy, n1i = Size_healthy, m2i = Mean_diseased,
    sd2i = SD_diseased, n2i = Size_diseased, data = data_study)
Sum_HD <- summary(diff_healthy_diseased)

diff_stable_exacerbated <- escalc(measure = "SMDH", m1i = Mean_stable,
    sd1i = SD_stable, n1i = Size_stable, m2i = Mean_exacerbated,
    sd2i = SD_exacerbated, n2i = Size_exacerbated, data = data_study)
Sum_SE <- summary(diff_stable_exacerbated)
```

Data preparation

```r
Simpson = c()
Variance = c()
Group = c()
Study = c()
Disease = c()
Sample = c()

coeff = which(is.na(diff_healthy_stable$yi) == FALSE)
name = rownames(diff_healthy_stable)[coeff]
Simpson = c(Simpson, diff_healthy_stable$yi[coeff])
Variance = c(Variance, diff_healthy_stable$vi[coeff])
Group = c(Group, rep("Healthy/Stable", length(diff_healthy_stable$yi[coeff])))
Study = c(Study, diff_healthy_stable$author[coeff])
Disease = c(Disease, diff_healthy_stable$Disease[coeff])
Sample = c(Sample, diff_healthy_stable$Sample[coeff])

coeff = which(is.na(diff_healthy_exacerbated$yi) == FALSE)
name = rownames(diff_healthy_exacerbated)[coeff]
Simpson = c(Simpson, diff_healthy_exacerbated$yi[coeff])
Variance = c(Variance, diff_healthy_exacerbated$vi[coeff])
Group = c(Group, rep("Healthy/Exacerbated",
↪   length(diff_healthy_exacerbated$yi[coeff])))
Study = c(Study, diff_healthy_exacerbated$author[coeff])
Disease = c(Disease, diff_healthy_exacerbated$Disease[coeff])
Sample = c(Sample, diff_healthy_exacerbated$Sample[coeff])

coeff = which(is.na(diff_healthy_diseased$yi) == FALSE)
name = rownames(diff_healthy_diseased)[coeff]
Simpson = c(Simpson, diff_healthy_diseased$yi[coeff])
Variance = c(Variance, diff_healthy_diseased$vi[coeff])
Group = c(Group, rep("Healthy/Diseased", length(diff_healthy_diseased$yi[coeff])))
Study = c(Study, diff_healthy_diseased$author[coeff])
Disease = c(Disease, diff_healthy_diseased$Disease[coeff])
Sample = c(Sample, diff_healthy_diseased$Sample[coeff])

coeff = which(is.na(diff_stable_exacerbated$yi) == FALSE)
name = rownames(diff_stable_exacerbated)[coeff]
Simpson = c(Simpson, diff_stable_exacerbated$yi[coeff])
Variance = c(Variance, diff_stable_exacerbated$vi[coeff])
Group = c(Group, rep("Stable/Exacerbated", length(diff_stable_exacerbated$yi[coeff])))
Study = c(Study, diff_stable_exacerbated$author[coeff])
Disease = c(Disease, diff_stable_exacerbated$Disease[coeff])
Sample = c(Sample, diff_stable_exacerbated$Sample[coeff])

data_ME <- data.frame(Simpson, Variance, Group, Study, Disease,
    Sample)
```

```r
data_ME$Simpson <- as.numeric(data_ME$Simpson)
data_ME$Variance <- as.numeric(data_ME$Variance)
data_ME$Study <- as.factor(data_ME$Study)
data_ME$Group <- as.factor(data_ME$Group)
data_ME$Group <- relevel(data_ME$Group, ref = "Healthy/Stable")
data_ME$Disease <- as.factor(data_ME$Disease)
data_ME$Sample <- as.factor(data_ME$Sample)

summary(data_ME)
```

Mixed effect model with metafor

```r
# WEIGHTED LINEAR MODEL FOR ALL SUMMARY
# ################################################################################
wlm0 <- lm(Simpson ~ 1, weights = 1/Variance, data = data_ME)
summary(wlm0)
wlm1 <- lm(Simpson ~ Group, weights = 1/Variance, data = data_ME)
summary(wlm1)
anova(wlm0, wlm1)

# Confidence interval
data_ME$Group <- relevel(data_ME$Group, ref = "Healthy/Stable")
HS <- rma(yi = Simpson, vi = Variance, mods = ~Group, method = "REML",
    data = data_ME)
summary(HS)
data_ME$Group <- relevel(data_ME$Group, ref = "Healthy/Exacerbated")
HE <- rma(yi = Simpson, vi = Variance, mods = ~Group, method = "REML",
    data = data_ME)
summary(HE)
data_ME$Group <- relevel(data_ME$Group, ref = "Healthy/Diseased")
HD <- rma(yi = Simpson, vi = Variance, mods = ~Group, method = "REML",
    data = data_ME)
summary(HD)
data_ME$Group <- relevel(data_ME$Group, ref = "Stable/Exacerbated")
SE <- rma(yi = Simpson, vi = Variance, mods = ~Group, method = "REML",
    data = data_ME)
summary(SE)

# WEIGHTED LINEAR MODEL FOR ASTHMA
# ################################################################################
coef_asthma <- which(data_ME$Disease == "Asthma")
data_study_asthma <- data_ME[coef_asthma, ]

wlm0 <- lm(Simpson ~ 1, weights = 1/Variance, data = data_study_asthma)
summary(wlm0)
wlm1 <- lm(Simpson ~ Group, weights = 1/Variance, data = data_study_asthma)
summary(wlm1)
anova(wlm0, wlm1)
```

```r
# Confidence interval
data_study_asthma$Group <- relevel(data_study_asthma$Group, ref = "Healthy/Diseased")
HD_asthma <- rma(yi = Simpson, vi = Variance, mods = ~Group,
    method = "REML", data = data_study_asthma)
summary(HD_asthma)
data_study_asthma$Group <- relevel(data_study_asthma$Group, ref = "Healthy/Stable")
HS_asthma <- rma(yi = Simpson, vi = Variance, mods = ~Group,
    method = "REML", data = data_study_asthma)
summary(HS_asthma)
data_study_asthma$Group <- relevel(data_study_asthma$Group, ref =
↪    "Healthy/Exacerbated")
HE_asthma <- rma(yi = Simpson, vi = Variance, mods = ~Group,
    method = "REML", data = data_study_asthma)
summary(HE_asthma)
data_study_asthma$Group <- relevel(data_study_asthma$Group, ref = "Stable/Exacerbated")
SE_asthma <- rma(yi = Simpson, vi = Variance, mods = ~Group,
    method = "REML", data = data_study_asthma)
summary(SE_asthma)

# WEIGHTED LINEAR MODEL FOR COPD
# ###############################################################################
coef_copd <- which(data_ME$Disease == "COPD")
data_study_copd <- data_ME[coef_copd, ]

wlm0 <- lm(Simpson ~ 1, weights = 1/Variance, data = data_study_copd)
summary(wlm0)
wlm1 <- lm(Simpson ~ Group, weights = 1/Variance, data = data_study_copd)
summary(wlm1)
anova(wlm0, wlm1)

# WEIGHTED LINEAR MODEL FOR CF
# ###############################################################################
coef_cf <- which(data_ME$Disease == "Cystic fibrosis")
data_study_cf <- data_ME[coef_cf, ]

wlm0 <- lm(Simpson ~ 1, weights = 1/Variance, data = data_study_cf)
summary(wlm0)
wlm1 <- lm(Simpson ~ Group, weights = 1/Variance, data = data_study_cf)
summary(wlm1)
anova(wlm0, wlm1)
```

Data preparation for the forestplot

```r
data_study_asthma_HS <- data_study_asthma
data_study_asthma_HS$Simpson[which(data_study_asthma_HS$Group !=
    "Healthy/Stable")] = NA
data_study_asthma_HS <- cbind(data_study_asthma_HS, c(Sum_HS$ci.lb[2],
```

```r
    rep(NA, 4)))
data_study_asthma_HS <- cbind(data_study_asthma_HS, c(Sum_HS$ci.ub[2],
    rep(NA, 4)))
colnames(data_study_asthma_HS)[7] <- "IC low"
colnames(data_study_asthma_HS)[8] <- "IC up"
data_study_copd_HS <- data_study_copd
data_study_copd_HS$Simpson[which(data_study_copd_HS$Group !=
    "Healthy/Stable")] = NA
data_study_copd_HS <- cbind(data_study_copd_HS, c(Sum_HS$ci.lb[5],
    NA))
data_study_copd_HS <- cbind(data_study_copd_HS, c(Sum_HS$ci.ub[5],
    NA))
colnames(data_study_copd_HS)[7] <- "IC low"
colnames(data_study_copd_HS)[8] <- "IC up"

data_study_asthma_HE <- data_study_asthma
data_study_asthma_HE$Simpson[which(data_study_asthma_HE$Group !=
    "Healthy/Exacerbated")] = NA
data_study_asthma_HE <- cbind(data_study_asthma_HE, c(NA, Sum_HE$ci.lb[2],
    rep(NA, 3)))
data_study_asthma_HE <- cbind(data_study_asthma_HE, c(NA, Sum_HE$ci.ub[2],
    rep(NA, 3)))
colnames(data_study_asthma_HE)[7] <- "IC low"
colnames(data_study_asthma_HE)[8] <- "IC up"
data_study_cf_HE <- data_study_cf
data_study_cf_HE$Simpson[which(data_study_cf_HE$Group != "Healthy/Exacerbated")] = NA
data_study_cf_HE <- cbind(data_study_cf_HE, c(Sum_HE$ci.lb[7],
    NA))
data_study_cf_HE <- cbind(data_study_cf_HE, c(Sum_HE$ci.ub[7],
    NA))
colnames(data_study_cf_HE)[7] <- "IC low"
colnames(data_study_cf_HE)[8] <- "IC up"

data_study_asthma_HD <- data_study_asthma
data_study_asthma_HD$Simpson[which(data_study_asthma_HD$Group !=
    "Healthy/Diseased")] = NA
data_study_asthma_HD <- cbind(data_study_asthma_HD, c(NA, NA,
    Sum_HD$ci.lb[1], Sum_HD$ci.lb[3], NA))
data_study_asthma_HD <- cbind(data_study_asthma_HD, c(NA, NA,
    Sum_HD$ci.ub[1], Sum_HD$ci.ub[3], NA))
colnames(data_study_asthma_HD)[7] <- "IC low"
colnames(data_study_asthma_HD)[8] <- "IC up"
data_study_copd_HD <- data_study_copd
data_study_copd_HD$Simpson[which(data_study_copd_HD$Group !=
    "Healthy/Diseased")] = NA
data_study_copd_HD <- cbind(data_study_copd_HD, c(NA, Sum_HD$ci.lb[4]))
data_study_copd_HD <- cbind(data_study_copd_HD, c(NA, Sum_HD$ci.ub[4]))
colnames(data_study_copd_HD)[7] <- "IC low"
colnames(data_study_copd_HD)[8] <- "IC up"
```

```r
data_study_asthma_SE <- data_study_asthma
data_study_asthma_SE$Simpson[which(data_study_asthma_SE$Group !=
    "Stable/Exacerbated")] = NA
data_study_asthma_SE <- cbind(data_study_asthma_SE, c(rep(NA,
    4), Sum_SE$ci.lb[2]))
data_study_asthma_SE <- cbind(data_study_asthma_SE, c(rep(NA,
    4), Sum_SE$ci.ub[2]))
colnames(data_study_asthma_SE)[7] <- "IC low"
colnames(data_study_asthma_SE)[8] <- "IC up"
data_study_cf_SE <- data_study_cf
data_study_cf_SE$Simpson[which(data_study_cf_SE$Group != "Stable/Exacerbated")] = NA
data_study_cf_SE <- cbind(data_study_cf_SE, c(NA, Sum_SE$ci.lb[6]))
data_study_cf_SE <- cbind(data_study_cf_SE, c(NA, Sum_SE$ci.ub[6]))
colnames(data_study_cf_SE)[7] <- "IC low"
colnames(data_study_cf_SE)[8] <- "IC up"

test_simpson <- data.frame(coef1 = c(NA, data_study_asthma_HS$Simpson[1],
    rep(NA, 2), HS_asthma$beta[1], NA, data_study_copd_HS$Simpson,
    NA, NA, NA, NA, NA, NA, HS$beta[1]), coef2 = c(NA, data_study_asthma_HE$Simpson[2],
    rep(NA, 2), HE_asthma$beta[1], NA, NA, NA, NA, data_study_cf_HE$Simpson,
    NA, NA, NA, HE$beta[1]), coef3 = c(NA, NA, data_study_asthma_HD$Simpson[3:4],
    HD_asthma$beta[1], NA, data_study_copd_HD$Simpson, NA, NA,
    NA, NA, NA, NA, HD$beta[1]), coef4 = c(NA, data_study_asthma_SE$Simpson[5],
    rep(NA, 2), SE_asthma$beta[1], NA, NA, NA, NA, data_study_cf_SE$Simpson,
    NA, data_ME$Simpson[which(data_ME$Disease == "Bronchiectasis")],
    NA, SE$beta[1]), low1 = c(NA, data_study_asthma_HS$`IC low`[1],
    rep(NA, 2), HS_asthma$ci.lb[1], NA, data_study_copd_HS$`IC low`,
    NA, NA, NA, NA, NA, NA, HS$ci.lb[1]), low2 = c(NA, data_study_asthma_HE$`IC
    low`[2],
    rep(NA, 2), HE_asthma$ci.lb[1], NA, NA, NA, NA, data_study_cf_HE$`IC low`,
    NA, NA, NA, HE$ci.lb[1]), low3 = c(NA, NA, data_study_asthma_HD$`IC low`[3:4],
    HD_asthma$ci.lb[1], NA, data_study_copd_HD$`IC low`, NA,
    NA, NA, NA, NA, NA, HD$ci.lb[1]), low4 = c(NA, data_study_asthma_SE$`IC low`[5],
    rep(NA, 2), SE_asthma$ci.lb[1], NA, NA, NA, NA, data_study_cf_SE$`IC low`,
    NA, Sum_SE$ci.lb[which(Sum_SE$Disease == "Bronchiectasis")],
    NA, SE$ci.lb[1]), high1 = c(NA, data_study_asthma_HS$`IC up`[1],
    rep(NA, 2), HS_asthma$ci.ub[1], NA, data_study_copd_HS$`IC up`,
    NA, NA, NA, NA, NA, NA, HS$ci.ub[1]), high2 = c(NA, data_study_asthma_HE$`IC
    up`[2],
    rep(NA, 2), HE_asthma$ci.ub[1], NA, NA, NA, NA, data_study_cf_HE$`IC up`,
    NA, NA, NA, HE$ci.ub[1]), high3 = c(NA, NA, data_study_asthma_HD$`IC up`[3:4],
    HD_asthma$ci.ub[1], NA, data_study_copd_HD$`IC up`, NA, NA,
    NA, NA, NA, NA, HD$ci.ub[1]), high4 = c(NA, data_study_asthma_SE$`IC up`[5],
    rep(NA, 2), SE_asthma$ci.ub[1], NA, NA, NA, NA, data_study_cf_SE$`IC up`,
    NA, Sum_SE$ci.ub[which(Sum_SE$Disease == "Bronchiectasis")],
    NA, SE$ci.ub[1]))

# Reorganisation of rows for classification by sample type
```

```r
test_simpson <- rbind(test_simpson[1, ], test_simpson[3, ], test_simpson[2,
    ], test_simpson[4:15, ])

# Execute the following line of code if you want the
# forestplot with all the diversity indexes before
# executing the file forest_plot_3.rmd test_simpson <-
#
↪   rbind(test_simpson[1,],NA,test_simpson[2,],NA,NA,NA,test_simpson[3,],NA,test_simpson[4,],NA,N
```

Forest plot

```r
col_no <- grep("coef", colnames(test_simpson))
row_names <- list(c("Disease", rep("Asthma", 3), "", "", rep("COPD",
    2), "", rep("CF", 2), "", "NCFB", "", ""), c("Sample", "BAL",
    "Sputum (induced)", "Sputum (induced)", "", "", "BAL", "LA",
    "", "Sputum", "Sputum", "", "BAL", "", ""), c("Study", "Denner et al. 2016",
    "Li et al. 2017", "Huang et al. 2020", "Asthma summary",
    "", "Pragman et al. 2012", "Kim et al. 2017", "", "Soret et al. 2020",
    "Filkins et al. 2012", "", "Byun et al. 2017", "", "Summary"))

coef <- with(test_simpson, cbind(coef1, coef3, coef4, coef2))
low <- with(test_simpson, cbind(low1, low3, low4, low2))
high <- with(test_simpson, cbind(high1, high3, high4, high2))
pdf(file = "difference_simpson_BC.pdf", width = 12, height = 10)
forestplot(row_names, coef, low, high, hrzl_lines = list(`2` = gpar(lty = 1),
    `6` = gpar(lty = 2), `9` = gpar(lty = 2), `12` = gpar(lty = 2),
    `14` = gpar(lty = 1)), title = "Difference of mean Simpson index",
    boxsize = 0.2, col = fpColors(box = c("royalblue", "orange",
        "green", "red"), line = c("royalblue", "orange", "green",
        "red"), summary = c("royalblue", "orange", "green", "red")),
    vertices = TRUE, is.summary = c(TRUE, rep(FALSE, 3), TRUE,
        rep(FALSE, 9), TRUE), xlab = "Difference of mean Simpson index",
    new_page = FALSE, legend = c("Healthy/Stable", "Healthy/Diseased",
        "Stable/Exacerbated", "Healthy/Exacerbated"), legend_args = fpLegend(pos =
        ↪   list(x = 0.85,
        y = 0.58), title = "Difference", r = unit(0.1, "snpc"),
        gp = gpar(col = "#CCCCCC", lwd = 1.5)))
dev.off()
```