

Chao1 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("Par14_1", "Den16", "Li17", "Gol13",
  "Lee18", "Hua20", "Par14_2", "Mil15", "Nar17", "Sor20")
```

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

```
Chao1 = c()
Variance = c()
Group = c()
Study = c()
Disease = c()
Sample = c()

coeff = which(is.na(diff_healthy_stable$yi) == FALSE)
name = rownames(diff_healthy_stable)[coeff]
Chao1 = c(Chao1, diff_healthy_stable$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$yi) == FALSE)
name = rownames(diff_healthy_exacerbated)[coeff]
Chao1 = c(Chao1, diff_healthy_exacerbated$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$yi) == FALSE)
name = rownames(diff_healthy_diseased)[coeff]
Chao1 = c(Chao1, diff_healthy_diseased$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$yi) == FALSE)
name = rownames(diff_stable_exacerbated)[coeff]
Chao1 = c(Chao1, diff_stable_exacerbated$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(Chao1, Variance, Group, Study, Disease,
  Sample)
```

```

data_ME$Chao1 <- as.numeric(data_ME$Chao1)
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

```

# WEIGHTED LINEAR MODEL FOR ALL SUMMARY
# #####

wlm0 <- lm(Chao1 ~ 1, weights = 1/Variance, data = data_ME)
summary(wlm0)
wlm1 <- lm(Chao1 ~ 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 = Chao1, vi = Variance, mods = ~Group, method = "REML",
  data = data_ME)
summary(HS)
data_ME$Group <- relevel(data_ME$Group, ref = "Healthy/Exacerbated")
HE <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "REML",
  data = data_ME)
summary(HE)
data_ME$Group <- relevel(data_ME$Group, ref = "Healthy/Diseased")
HD <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "REML",
  data = data_ME)
summary(HD)
data_ME$Group <- relevel(data_ME$Group, ref = "Stable/Exacerbated")
SE <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "REML",
  data = data_ME)
summary(SE)

# WEIGHTED LINEAR MODEL FOR ASTHMA
# #####

# All samples #####
coef_asthma <- which(data_ME$Disease == "Asthma")
data_study_asthma <- data_ME[coef_asthma, ]

wlm0 <- lm(Chao1 ~ 1, weights = 1/Variance, data = data_study_asthma)
summary(wlm0)
wlm1 <- lm(Chao1 ~ Group, weights = 1/Variance, data = data_study_asthma)

```

```

summary(wlm1)
anova(wlm0, wlm1)

# Confidence interval
data_study_asthma$Group <- relevel(data_study_asthma$Group, ref = "Healthy/Stable")
HS_asthma <- rma(yi = Chao1, 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 = Chao1, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_asthma)
summary(HE_asthma)
data_study_asthma$Group <- relevel(data_study_asthma$Group, ref = "Healthy/Diseased")
HD_asthma <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_asthma)
summary(HD_asthma)
data_study_asthma$Group <- relevel(data_study_asthma$Group, ref = "Stable/Exacerbated")
SE_asthma <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_asthma)
summary(SE_asthma)

# UA #####
coef_asthma_ua <- which(data_study_asthma$Sample == "UA")
data_study_asthma_ua <- data_study_asthma[coef_asthma_ua, ]

# wlm0<-lm(Chao1~1, weights=1/Variance,
# data=data_study_asthma_ua) summary(wlm0)
# wlm1<-lm(Chao1~Group,
# weights=1/Variance,data=data_study_asthma_ua)
# summary(wlm1) anova(wlm0,wlm1)

# Confidence interval
data_study_asthma_ua$Group <- relevel(data_study_asthma_ua$Group,
  ref = "Healthy/Stable")
HS_asthma_ua <- rma(yi = Chao1, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma_ua)
summary(HS_asthma_ua)

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

wlm0 <- lm(Chao1 ~ 1, weights = 1/Variance, data = data_study_copd)
summary(wlm0)
wlm1 <- lm(Chao1 ~ Group, weights = 1/Variance, data = data_study_copd)
summary(wlm1)
anova(wlm0, wlm1)

```

```

# Confidence interval
data_study_copd$Group <- relevel(data_study_copd$Group, ref = "Healthy/Stable")
HS_copd <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "FE",
  data = data_study_copd)
summary(HS_copd)
data_study_copd$Group <- relevel(data_study_copd$Group, ref = "Stable/Exacerbated")
SE_copd <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "FE",
  data = data_study_copd)
summary(SE_copd)

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

wlm0 <- lm(Chao1 ~ 1, weights = 1/Variance, data = data_study_cf)
summary(wlm0)
wlm1 <- lm(Chao1 ~ Group, weights = 1/Variance, data = data_study_cf)
summary(wlm1)
anova(wlm0, wlm1)

# Confidence interval
data_study_copd$Group <- relevel(data_study_copd$Group, ref = "Healthy/Diseased")
HD_cf <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "FE",
  data = data_study_copd)
summary(HD_cf)

data_study_copd$Group <- relevel(data_study_copd$Group, ref = "Healthy/Exacerbated")
HE_cf <- rma(yi = Chao1, vi = Variance, mods = ~Group, method = "FE",
  data = data_study_copd)
summary(HE_cf)

```

Data preparation for the forestplot

```

data_study_asthma_HS <- data_study_asthma
data_study_asthma_HS$Chao1[which(data_study_asthma_HS$Group !=
  "Healthy/Stable")] = NA
data_study_asthma_HS <- cbind(data_study_asthma_HS, c(Sum_HS$ci.lb[1],
  Sum_HS$ci.lb[3:5], rep(NA, 4)))
data_study_asthma_HS <- cbind(data_study_asthma_HS, c(Sum_HS$ci.ub[1],
  Sum_HS$ci.ub[3:5], 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$Chao1[which(data_study_copd_HS$Group != "Healthy/Stable")] = NA
data_study_copd_HS <- cbind(data_study_copd_HS, c(Sum_HS$ci.lb[7],
  NA))

```

```

data_study_copd_HS <- cbind(data_study_copd_HS, c(Sum_HS$ci.ub[7],
  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$Chao1[which(data_study_asthma_HE$Group !=
  "Healthy/Exacerbated")] = NA
data_study_asthma_HE <- cbind(data_study_asthma_HE, c(rep(NA,
  4), Sum_HE$ci.lb[3], rep(NA, 3)))
data_study_asthma_HE <- cbind(data_study_asthma_HE, c(rep(NA,
  4), Sum_HE$ci.ub[3], 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$Chao1[which(data_study_cf_HE$Group != "Healthy/Exacerbated")] = NA
data_study_cf_HE <- cbind(data_study_cf_HE, c(Sum_HE$ci.lb[10],
  NA))
data_study_cf_HE <- cbind(data_study_cf_HE, c(Sum_HE$ci.ub[10],
  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$Chao1[which(data_study_asthma_HD$Group !=
  "Healthy/Diseased")] = NA
data_study_asthma_HD <- cbind(data_study_asthma_HD, c(rep(NA,
  5), Sum_HD$ci.lb[2], Sum_HD$ci.lb[6], NA))
data_study_asthma_HD <- cbind(data_study_asthma_HD, c(rep(NA,
  5), Sum_HD$ci.ub[2], Sum_HD$ci.ub[6], NA))
colnames(data_study_asthma_HD)[7] <- "IC low"
colnames(data_study_asthma_HD)[8] <- "IC up"
data_study_cf_HD <- data_study_cf
data_study_cf_HD$Chao1[which(data_study_cf_HD$Group != "Healthy/Diseased")] = NA
data_study_cf_HD <- cbind(data_study_cf_HD, c(NA, Sum_HD$ci.lb[9]))
data_study_cf_HD <- cbind(data_study_cf_HD, c(NA, Sum_HD$ci.ub[9]))
colnames(data_study_cf_HD)[7] <- "IC low"
colnames(data_study_cf_HD)[8] <- "IC up"

data_study_asthma_SE <- data_study_asthma
data_study_asthma_SE$Chao1[which(data_study_asthma_SE$Group !=
  "Stable/Exacerbated")] = NA
data_study_asthma_SE <- cbind(data_study_asthma_SE, c(rep(NA,
  7), Sum_SE$ci.lb[3]))
data_study_asthma_SE <- cbind(data_study_asthma_SE, c(rep(NA,
  7), Sum_SE$ci.ub[3]))
colnames(data_study_asthma_SE)[7] <- "IC low"
colnames(data_study_asthma_SE)[8] <- "IC up"
data_study_copd_SE <- data_study_copd

```

```

data_study_copd_SE$Chao1[which(data_study_copd_SE$Group != "Stable/Exacerbated")] = NA
data_study_copd_SE <- cbind(data_study_copd_SE, c(NA, Sum_SE$ci.lb[8]))
data_study_copd_SE <- cbind(data_study_copd_SE, c(NA, Sum_SE$ci.ub[8]))
colnames(data_study_copd_SE)[7] <- "IC low"
colnames(data_study_copd_SE)[8] <- "IC up"

test_chao1 <- data.frame(coef1 = c(NA, data_study_asthma_HS$Chao1[1:4],
  rep(NA, 2), HS_asthma$beta[1], NA, data_study_copd_HS$Chao1,
  NA, NA, NA, NA, HS$beta[1]), coef2 = c(NA, NA, data_study_asthma_HE$Chao1[5],
  rep(NA, 4), HE_asthma$beta[1], NA, NA, NA, NA, data_study_cf_HE$Chao1,
  NA, HE$beta[1]), coef3 = c(NA, rep(NA, 4), data_study_asthma_HD$Chao1[6:7],
  HD_asthma$beta[1], NA, NA, NA, NA, NA, data_study_cf_HD$Chao1[2],
  NA, HD$beta[1]), coef4 = c(NA, NA, data_study_asthma_SE$Chao1[8],
  rep(NA, 4), SE_asthma$beta[1], NA, NA, data_study_copd_SE$Chao1[2],
  NA, NA, NA, NA, SE$beta[1]), low1 = c(NA, data_study_asthma_HS$`IC low`[1:4],
  rep(NA, 2), HS_asthma$ci.lb[1], NA, data_study_copd_HS$`IC low`,
  NA, NA, NA, NA, HS$ci.lb[1]), low2 = c(NA, NA, data_study_asthma_HE$`IC low`[5],
  rep(NA, 4), HE_asthma$ci.lb[1], NA, NA, NA, NA, data_study_cf_HE$`IC low`,
  NA, HE$ci.lb[1]), low3 = c(NA, rep(NA, 4), data_study_asthma_HD$`IC low`[6:7],
  HD_asthma$ci.lb[1], NA, NA, NA, NA, NA, data_study_cf_HD$`IC low`[2],
  NA, HD$ci.lb[1]), low4 = c(NA, NA, data_study_asthma_SE$`IC low`[8],
  rep(NA, 4), SE_asthma$ci.lb[1], NA, NA, data_study_copd_SE$`IC low`[2],
  NA, NA, NA, NA, SE$ci.lb[1]), high1 = c(NA, data_study_asthma_HS$`IC up`[1:4],
  rep(NA, 2), HS_asthma$ci.ub[1], NA, data_study_copd_HS$`IC up`,
  NA, NA, NA, NA, HS$ci.ub[1]), high2 = c(NA, NA, data_study_asthma_HE$`IC up`[5],
  rep(NA, 4), HE_asthma$ci.ub[1], NA, NA, NA, NA, data_study_cf_HE$`IC up`,
  NA, HE$ci.ub[1]), high3 = c(NA, rep(NA, 4), data_study_asthma_HD$`IC up`[6:7],
  HD_asthma$ci.ub[1], NA, NA, NA, NA, NA, data_study_cf_HD$`IC up`[2],
  NA, HD$ci.ub[1]), high4 = c(NA, NA, data_study_asthma_SE$`IC up`[8],
  rep(NA, 4), SE_asthma$ci.ub[1], NA, NA, data_study_copd_SE$`IC up`[2],
  NA, NA, NA, NA, SE$ci.ub[1]))

inter <- test_chao1[13, ]
test_chao1[13, ] <- test_chao1[14, ]
test_chao1[14, ] <- inter

# Reorganisation of rows for classification by sample type
test_chao1 <- rbind(test_chao1[1, ], test_chao1[4, ], test_chao1[6,
  ], test_chao1[3, ], test_chao1[7, ], test_chao1[2, ], test_chao1[5,
  ], test_chao1[8:9, ], test_chao1[11, ], test_chao1[10, ],
  test_chao1[12:16, ])

# Addition of intermediate summaries for sample types
asthma_ua <- c(HS_asthma_ua$beta[1], NA, NA, NA, HS_asthma_ua$ci.lb[1],
  NA, NA, NA, HS_asthma_ua$ci.ub[1], NA, NA, NA)

test_chao1 <- rbind(test_chao1[1:7, ], asthma_ua, test_chao1[8:16,
  ])

```

```

# 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_chao1 <-
#
↪ rbind(test_chao1[1:3,],NA,NA,NA,test_chao1[4,],NA,test_chao1[5,],NA,NA,test_chao1[6:10,],NA,NA,NA)

```

Forest plot

```

col_no <- grep("coef", colnames(test_chao1))
row_names <- list(c("Disease", rep("Asthma", 6), "", "", "",
  rep("COPD", 2), "", rep("CF", 2), "", ""), c("Sample", "BAL",
  "BAL", "Sputum (induced)", "Sputum (induced)", "UA", "UA",
  "UA summary", "", "", "Sputum", "UA", "", "Sputum", "Sputum",
  "", ""), c("Study", "Goleva et al. 2013", "Denner et al. 2016",
  "Li et al. 2017", "Huang et al. 2020", "Park et al. 2014",
  "Lee et al. 2018", "", "Asthma summary", "", "Millares et al. 2015",
  "Park et al. 2014", "", "Narayanamurthy et al. 2017", "Soret et al. 2020",
  "", "Summary"))

coef <- with(test_chao1, cbind(coef1, coef3, coef4, coef2))
low <- with(test_chao1, cbind(low1, low3, low4, low2))
high <- with(test_chao1, cbind(high1, high3, high4, high2))
pdf(file = "difference_chao1_BC.pdf", width = 12, height = 10)
forestplot(row_names, coef, low, high, hrzl_lines = list(`2` = gpar(lty = 1),
  `10` = gpar(lty = 2), `13` = gpar(lty = 2), `16` = gpar(lty = 1)),
  title = "Difference of mean Chao1 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, 6), TRUE, TRUE, rep(FALSE, 7), TRUE), xlab = "Difference of mean
    ↪ Chao1 index",
  new_page = FALSE, legend = c("Healthy/Stable", "Healthy/Diseased",
    "Stable/Exacerbated", "Healthy/Exacerbated"), legend_args = fpLegend(pos =
    ↪ list(x = 0.87,
    y = 0.85), title = "Difference", r = unit(0.1, "snpc"),
    gp = gpar(col = "#CCCCCC", lwd = 1.5)))
dev.off()

```