

Shannon 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",
  "Liu20", "Mun16", "Sve17", "Mar13", "Lee18", "Hua20", "Par14_2",
  "Ein16", "Pra12", "Mil15", "Fei17", "Erb11", "Kim17", "Wan20",
  "Ple19", "Nar17", "Cob15", "Car13", "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

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

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

data_ME$Shannon <- as.numeric(data_ME$Shannon)
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 GLOBAL INTER-DISEASE SUMMARY
# #####
wlm0 <- lm(Shannon ~ 1, weights = 1/Variance, data = data_ME)
summary(wlm0)
wlm1 <- lm(Shannon ~ 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 = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_ME)
summary(HS)
data_ME$Group <- relevel(data_ME$Group, ref = "Healthy/Exacerbated")
HE <- rma(yi = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_ME)
summary(HE)
data_ME$Group <- relevel(data_ME$Group, ref = "Healthy/Diseased")
HD <- rma(yi = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_ME)
summary(HD)
data_ME$Group <- relevel(data_ME$Group, ref = "Stable/Exacerbated")
SE <- rma(yi = Shannon, 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(Shannon ~ 1, weights = 1/Variance, data = data_study_asthma)

```

```

summary(wlm0)
wlm1 <- lm(Shannon ~ 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 = Shannon, 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 = Shannon, 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 = Shannon, 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 = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma)
summary(SE_asthma)

# Sputum (induced) #####
coef_asthma_si <- which(data_study_asthma$Sample == "Sputum (induced)")
data_study_asthma_si <- data_study_asthma[coef_asthma_si, ]

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

# Confidence interval
data_study_asthma_si$Group <- relevel(data_study_asthma_si$Group,
  ref = "Healthy/Stable")
HS_asthma_si <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma_si)
summary(HS_asthma_si)
data_study_asthma_si$Group <- relevel(data_study_asthma_si$Group,
  ref = "Healthy/Exacerbated")
HE_asthma_si <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma_si)
summary(HE_asthma_si)
data_study_asthma_si$Group <- relevel(data_study_asthma_si$Group,
  ref = "Healthy/Diseased")
HD_asthma_si <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma_si)

```

```

summary(HD_asthma_si)
data_study_asthma_si$Group <- relevel(data_study_asthma_si$Group,
  ref = "Stable/Exacerbated")
SE_asthma_si <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma_si)
summary(SE_asthma_si)

# BAL #####
coef_asthma_bal <- which(data_study_asthma$Sample == "BAL")
data_study_asthma_bal <- data_study_asthma[coef_asthma_bal, ]

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

# Confidence interval
data_study_asthma_bal$Group <- relevel(data_study_asthma_bal$Group,
  ref = "Healthy/Stable")
HS_asthma_bal <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma_bal)
summary(HS_asthma_bal)
data_study_asthma_bal$Group <- relevel(data_study_asthma_bal$Group,
  ref = "Healthy/Diseased")
HD_asthma_bal <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma_bal)
summary(HD_asthma_bal)

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

# wlm0<-lm(Shannon~1, weights=1/Variance,
# data=data_study_asthma_ua) summary(wlm0)
# wlm1<-lm(Shannon~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 = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_asthma_ua)
summary(HS_asthma_ua)

# WEIGHTED LINEAR MODEL FOR COPD
# #####

```

```

# All samples #####
coef_copd <- which(data_ME$Disease == "COPD")
data_study_copd <- data_ME[coef_copd, ]

wlm0 <- lm(Shannon ~ 1, weights = 1/Variance, data = data_study_copd)
summary(wlm0)
wlm1 <- lm(Shannon ~ 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 = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_copd)
summary(HS_copd)
data_study_copd$Group <- relevel(data_study_copd$Group, ref = "Healthy/Exacerbated")
HE_copd <- rma(yi = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_copd)
summary(HE_copd)
data_study_copd$Group <- relevel(data_study_copd$Group, ref = "Healthy/Diseased")
HD_copd <- rma(yi = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_copd)
summary(HD_copd)
data_study_copd$Group <- relevel(data_study_copd$Group, ref = "Stable/Exacerbated")
SE_copd <- rma(yi = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_copd)
summary(SE_copd)

# BAL #####
coef_copd_bal <- which(data_study_copd$Sample == "BAL")
data_study_copd_bal <- data_study_copd[coef_copd_bal, ]

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

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

# WEIGHTED LINEAR MODEL FOR CF
#####
# All samples #####
coef_cf <- which(data_ME$Disease == "Cystic fibrosis")

```

```

data_study_cf <- data_ME[coef_cf, ]

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

# Confidence interval
data_study_cf$Group <- relevel(data_study_cf$Group, ref = "Healthy/Exacerbated")
HE_cf <- rma(yi = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_cf)
summary(HE_cf)
data_study_cf$Group <- relevel(data_study_cf$Group, ref = "Healthy/Diseased")
HD_cf <- rma(yi = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_cf)
summary(HD_cf)
data_study_cf$Group <- relevel(data_study_cf$Group, ref = "Stable/Exacerbated")
SE_cf <- rma(yi = Shannon, vi = Variance, mods = ~Group, method = "REML",
  data = data_study_cf)
summary(SE_cf)

# Sputum #####
coef_cf_sputum <- which(data_study_cf$Sample == "Sputum")
data_study_cf_sputum <- data_study_cf[coef_cf_sputum, ]

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

# Confidence interval
data_study_cf_sputum$Group <- relevel(data_study_cf_sputum$Group,
  ref = "Healthy/Exacerbated")
HE_cf_sputum <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_cf_sputum)
summary(HE_cf_sputum)
data_study_cf_sputum$Group <- relevel(data_study_cf_sputum$Group,
  ref = "Healthy/Diseased")
HD_cf_sputum <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_cf_sputum)
summary(HD_cf_sputum)
data_study_cf_sputum$Group <- relevel(data_study_cf_sputum$Group,
  ref = "Stable/Exacerbated")
SE_cf_sputum <- rma(yi = Shannon, vi = Variance, mods = ~Group,
  method = "REML", data = data_study_cf_sputum)
summary(SE_cf_sputum)

```

Data preparation for the forestplot

```
data_study_asthma_HS <- data_study_asthma
data_study_asthma_HS$Shannon[which(data_study_asthma_HS$Group !=
  "Healthy/Stable")] = NA
data_study_asthma_HS <- cbind(data_study_asthma_HS, c(Sum_HS$ci.lb[which(Sum_HS$Disease
  ==
  "Asthma")][1], Sum_HS$ci.lb[which(Sum_HS$Disease == "Asthma")][3:5],
  Sum_HS$ci.lb[which(Sum_HS$Disease == "Asthma")][8:9], rep(NA,
    6)))
data_study_asthma_HS <- cbind(data_study_asthma_HS, c(Sum_HS$ci.ub[which(Sum_HS$Disease
  ==
  "Asthma")][1], Sum_HS$ci.ub[which(Sum_HS$Disease == "Asthma")][3:5],
  Sum_HS$ci.ub[which(Sum_HS$Disease == "Asthma")][8:9], rep(NA,
    6)))
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$Shannon[which(data_study_copd_HS$Group !=
  "Healthy/Stable")] = NA
data_study_copd_HS <- cbind(data_study_copd_HS, c(Sum_HS$ci.lb[which(Sum_HS$Disease ==
  "COPD")][1:3], Sum_HS$ci.lb[which(Sum_HS$Disease == "COPD")][6],
  Sum_HS$ci.lb[which(Sum_HS$Disease == "COPD")][8], rep(NA,
    3)))
data_study_copd_HS <- cbind(data_study_copd_HS, c(Sum_HS$ci.ub[which(Sum_HS$Disease ==
  "COPD")][1:3], Sum_HS$ci.ub[which(Sum_HS$Disease == "COPD")][6],
  Sum_HS$ci.ub[which(Sum_HS$Disease == "COPD")][8], rep(NA,
    3)))
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$Shannon[which(data_study_asthma_HE$Group !=
  "Healthy/Exacerbated")] = NA
data_study_asthma_HE <- cbind(data_study_asthma_HE, c(rep(NA,
  6), Sum_HE$ci.lb[which(Sum_HE$Disease == "Asthma")][3], rep(NA,
  5)))
data_study_asthma_HE <- cbind(data_study_asthma_HE, c(rep(NA,
  6), Sum_HE$ci.ub[which(Sum_HE$Disease == "Asthma")][3], rep(NA,
  5)))
colnames(data_study_asthma_HE)[7] <- "IC low"
colnames(data_study_asthma_HE)[8] <- "IC up"
data_study_copd_HE <- data_study_copd
data_study_copd_HE$Shannon[which(data_study_copd_HE$Group !=
  "Healthy/Exacerbated")] = NA
data_study_copd_HE <- cbind(data_study_copd_HE, c(rep(NA, 8)))
data_study_copd_HE <- cbind(data_study_copd_HE, c(rep(NA, 8)))
colnames(data_study_copd_HE)[7] <- "IC low"
colnames(data_study_copd_HE)[8] <- "IC up"
```



```

data_study_cf_HE <- data_study_cf
data_study_cf_HE$Shannon[which(data_study_cf_HE$Group != "Healthy/Exacerbated")] = NA
data_study_cf_HE <- cbind(data_study_cf_HE, c(Sum_HE$ci.lb[which(Sum_HE$Disease ==
  "Cystic fibrosis")][5], rep(NA, 4)))
data_study_cf_HE <- cbind(data_study_cf_HE, c(Sum_HE$ci.ub[which(Sum_HE$Disease ==
  "Cystic fibrosis")][5], rep(NA, 4)))
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$Shannon[which(data_study_asthma_HD$Group !=
  "Healthy/Diseased")] = NA
data_study_asthma_HD <- cbind(data_study_asthma_HD, c(rep(NA,
  7), Sum_HD$ci.lb[which(Sum_HD$Disease == "Asthma")][2],
  ↪ Sum_HD$ci.lb[which(Sum_HD$Disease ==
  "Asthma")][6:7], Sum_HD$ci.lb[which(Sum_HD$Disease == "Asthma")][10],
  NA))
data_study_asthma_HD <- cbind(data_study_asthma_HD, c(rep(NA,
  7), Sum_HD$ci.ub[which(Sum_HD$Disease == "Asthma")][2],
  ↪ Sum_HD$ci.ub[which(Sum_HD$Disease ==
  "Asthma")][6:7], Sum_HD$ci.ub[which(Sum_HD$Disease == "Asthma")][10],
  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$Shannon[which(data_study_copd_HD$Group !=
  "Healthy/Diseased")] = NA
data_study_copd_HD <- cbind(data_study_copd_HD, c(rep(NA, 5),
  Sum_HD$ci.lb[which(Sum_HD$Disease == "COPD")][5], Sum_HD$ci.lb[which(Sum_HD$Disease
  ↪ ==
  "COPD")][7], NA))
data_study_copd_HD <- cbind(data_study_copd_HD, c(rep(NA, 5),
  Sum_HD$ci.ub[which(Sum_HD$Disease == "COPD")][5], Sum_HD$ci.ub[which(Sum_HD$Disease
  ↪ ==
  "COPD")][7], NA))
colnames(data_study_copd_HD)[7] <- "IC low"
colnames(data_study_copd_HD)[8] <- "IC up"
data_study_cf_HD <- data_study_cf
data_study_cf_HD$Shannon[which(data_study_cf_HD$Group != "Healthy/Diseased")] = NA
data_study_cf_HD <- cbind(data_study_cf_HD, c(rep(NA, 1),
  ↪ Sum_HD$ci.lb[which(Sum_HD$Disease ==
  "Cystic fibrosis")][1:2], rep(NA, 2)))
data_study_cf_HD <- cbind(data_study_cf_HD, c(rep(NA, 1),
  ↪ Sum_HD$ci.ub[which(Sum_HD$Disease ==
  "Cystic fibrosis")][1:2], rep(NA, 2)))
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$Shannon[which(data_study_asthma_SE$Group !=
  "Stable/Exacerbated")] = NA
data_study_asthma_SE <- cbind(data_study_asthma_SE, c(rep(NA,
  11), Sum_SE$ci.lb[which(Sum_SE$Disease == "Asthma")][3]))
data_study_asthma_SE <- cbind(data_study_asthma_SE, c(rep(NA,
  11), Sum_SE$ci.ub[which(Sum_SE$Disease == "Asthma")][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$Shannon[which(data_study_copd_SE$Group !=
  "Stable/Exacerbated")] = NA
data_study_copd_SE <- cbind(data_study_copd_SE, c(rep(NA, 7),
  Sum_SE$ci.lb[which(Sum_SE$Disease == "COPD")][4]))
data_study_copd_SE <- cbind(data_study_copd_SE, c(rep(NA, 7),
  Sum_SE$ci.ub[which(Sum_SE$Disease == "COPD")][4]))
colnames(data_study_copd_SE)[7] <- "IC low"
colnames(data_study_copd_SE)[8] <- "IC up"
data_study_cf_SE <- data_study_cf
data_study_cf_SE$Shannon[which(data_study_cf_SE$Group != "Stable/Exacerbated")] = NA
data_study_cf_SE <- cbind(data_study_cf_SE, c(rep(NA, 3),
  ↪ Sum_SE$ci.lb[which(Sum_SE$Disease ==
    "Cystic fibrosis")][3:4]))
data_study_cf_SE <- cbind(data_study_cf_SE, c(rep(NA, 3),
  ↪ Sum_SE$ci.ub[which(Sum_SE$Disease ==
    "Cystic fibrosis")][3:4]))
colnames(data_study_cf_SE)[7] <- "IC low"
colnames(data_study_cf_SE)[8] <- "IC up"

test_shannon <- data.frame(coef1 = c(NA, data_study_asthma_HS$Shannon[1:6],
  rep(NA, 4), HS_asthma$beta[1], NA, data_study_copd_HS$Shannon[1:5],
  rep(NA, 3), HS_copd$beta[1], NA, rep(NA, 9), HS$beta[1]),
  coef2 = c(NA, NA, data_study_asthma_HE$Shannon[which(data_study_asthma_HE$Shannon
    ↪ !=
      "NA")], rep(NA, 8), HE_asthma$beta[1], NA, rep(NA, 8),
      NA, NA, data_study_cf_HE$Shannon, HE_cf$beta[1], NA,
      NA, NA, HE$beta[1]), coef3 = c(NA, rep(NA, 6),
      ↪ data_study_asthma_HD$Shannon[8:11],
      HD_asthma$beta[1], NA, data_study_copd_HD$Shannon, HD_copd$beta[1],
      NA, data_study_cf_HD$Shannon, HD_cf$beta[1], NA, NA,
      NA, HD$beta[1]), coef4 = c(NA, NA, data_study_asthma_SE$Shannon[12],
      rep(NA, 8), SE_asthma$beta[1], NA, data_study_copd_SE$Shannon,
      SE_copd$beta[1], NA, data_study_cf_SE$Shannon, SE_cf$beta[1],
      NA, data_ME$Shannon[which(data_ME$Disease == "Bronchiectasis")],
      NA, SE$beta[1]), low1 = c(NA, data_study_asthma_HS$`IC low`[1:6],
      rep(NA, 4), HS_asthma$ci.lb[1], NA, data_study_copd_HS$`IC low`[1:5],
      rep(NA, 3), HS_copd$ci.lb[1], NA, rep(NA, 9), HS$ci.lb[1]),
      low2 = c(NA, NA, data_study_asthma_HE$`IC low`[which(data_study_asthma_HE$Shannon
        ↪ !=
          "NA")], rep(NA, 8), HE_asthma$ci.lb[1], NA, rep(NA, 8),

```

```

NA, NA, data_study_cf_HE$`IC low`, HE_cf$ci.lb[1], NA,
NA, NA, HE$ci.lb[1]), low3 = c(NA, rep(NA, 6), data_study_asthma_HD$`IC
↪ low`[8:11],
HD_asthma$ci.lb[1], NA, data_study_copd_HD$`IC low`,
HD_copd$ci.lb[1], NA, data_study_cf_HD$`IC low`, HD_cf$ci.lb[1],
NA, NA, NA, HD$ci.lb[1]), low4 = c(NA, NA, data_study_asthma_SE$`IC low`[12],
rep(NA, 8), SE_asthma$ci.lb[1], NA, data_study_copd_SE$`IC low`,
SE_copd$ci.lb[1], NA, data_study_cf_SE$`IC low`, SE_cf$ci.lb[1],
NA, Sum_SE$ci.lb[which(Sum_SE$Disease == "Bronchiectasis")],
NA, SE$ci.lb[1]), high1 = c(NA, data_study_asthma_HS$`IC up`[1:6],
rep(NA, 4), HS_asthma$ci.ub[1], NA, data_study_copd_HS$`IC up`[1:5],
rep(NA, 3), HS_copd$ci.ub[1], NA, rep(NA, 9), HS$ci.ub[1]),
high2 = c(NA, NA, data_study_asthma_HE$`IC up`[which(data_study_asthma_HE$Shannon
↪ !=
"NA")], rep(NA, 8), HE_asthma$ci.ub[1], NA, rep(NA, 8),
NA, NA, data_study_cf_HE$`IC up`, HE_cf$ci.ub[1], NA,
NA, NA, HE$ci.ub[1]), high3 = c(NA, rep(NA, 6), data_study_asthma_HD$`IC
↪ up`[8:11],
HD_asthma$ci.ub[1], NA, data_study_copd_HD$`IC up`, HD_copd$ci.ub[1],
NA, data_study_cf_HD$`IC up`, HD_cf$ci.ub[1], NA, NA,
NA, HD$ci.ub[1]), high4 = c(NA, NA, data_study_asthma_SE$`IC up`[12],
rep(NA, 8), SE_asthma$ci.ub[1], NA, data_study_copd_SE$`IC up`,
SE_copd$ci.ub[1], NA, data_study_cf_SE$`IC up`, SE_cf$ci.ub[1],
NA, Sum_SE$ci.ub[which(Sum_SE$Disease == "Bronchiectasis")],
NA, SE$ci.ub[1]))

```

Reorganisation of rows for classification by sample type

```

test_shannon <- rbind(test_shannon[1, ], test_shannon[4, ], test_shannon[8,
], test_shannon[10, ], test_shannon[5, ], test_shannon[3,
], test_shannon[6, ], test_shannon[11, ], test_shannon[9,
], test_shannon[2, ], test_shannon[7, ], test_shannon[12:13,
], test_shannon[17, ], test_shannon[16, ], test_shannon[15,
], test_shannon[20, ], test_shannon[19, ], test_shannon[21,
], test_shannon[18, ], test_shannon[14, ], test_shannon[22:23,
], test_shannon[25, ], test_shannon[24, ], test_shannon[26:33,
])

```

Addition of intermediate summaries for sample types

```

asthma_sputum_induced <- c(HS_asthma_si$beta[1], HE_asthma_si$beta[1],
HD_asthma_si$beta[1], SE_asthma_si$beta[1], HS_asthma_si$ci.lb[1],
HE_asthma_si$ci.lb[1], HD_asthma_si$ci.lb[1], SE_asthma_si$ci.lb[1],
HS_asthma_si$ci.ub[1], HE_asthma_si$ci.ub[1], HD_asthma_si$ci.ub[1],
SE_asthma_si$ci.ub[1])

```

```

asthma_bal <- c(HS_asthma_bal$beta[1], NA, HD_asthma_bal$beta[1],
NA, HS_asthma_bal$ci.lb[1], NA, HD_asthma_bal$ci.lb[1], NA,
HS_asthma_bal$ci.ub[1], NA, HD_asthma_bal$ci.ub[1], NA)

```

```

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)

copd_bal <- c(HS_copd_bal$beta[1], NA, NA, NA, HS_copd_bal$ci.lb[1],
  NA, NA, NA, HS_copd_bal$ci.ub[1], NA, NA, NA)

cf_sputum <- c(NA, HE_cf_sputum$beta[1], HD_cf_sputum$beta[1],
  SE_cf_sputum$beta[1], NA, HE_cf_sputum$ci.lb[1], HD_cf_sputum$ci.lb[1],
  SE_cf_sputum$ci.lb[1], NA, HE_cf_sputum$ci.ub[1], HD_cf_sputum$ci.ub[1],
  SE_cf_sputum$ci.ub[1])

test_shannon <- rbind(test_shannon[1:4, ], asthma_bal, test_shannon[5:9,
  ], asthma_sputum_induced, test_shannon[10:11, ], asthma_ua,
  test_shannon[12:15, ], copd_bal, test_shannon[16:28, ], cf_sputum,
  test_shannon[29:33, ])

# 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_shannon <-
# rbind(test_shannon[1:30,],NA,test_shannon[31:38,])

```

Forest plot

```

col_no <- grep("coef", colnames(test_shannon))
row_names <- list(c("Disease", rep("Asthma", 3), "", rep("Asthma",
  5), "", rep("Asthma", 2), "", "", "", rep("COPD", 2), "",
  rep("COPD", 6), "", "", rep("CF", 5), "", "", "", "NCFB",
  "", ""), c("Sample", rep("BAL", 3), "BAL summary", "Sputum",
  rep("Sputum (induced)", 4), "Sputum (induced) summary", "UA",
  "UA", "UA summary", "", "", rep("BAL", 2), "BAL summary",
  "LA", "LA", "Sputum", "Sputum", "Sputum (induced)", "UA",
  "", "", "LA", rep("Sputum", 4), "Sputum summary", "", "",
  "BAL", "", "")), c("Study", "Goleva et al. 2013", "Denner et al. 2016",
  "Sverrild et al. 2017", "", "Liu et al. 2020", "Li et al. 2017",
  "Marri et al. 2013", "Huang et al. 2020", "Munck et al. 2016",
  "", "Park et al. 2014", "Lee et al. 2018", "", "Asthma summary",
  "", "Erb-Downward et al. 2011", "Pragman et al. 2012", "",
  "Einarsson et al. 2016", "Kim et al. 2017", "Feigelman et al. 2017",
  "Millares et al. 2015", "Wang et al. 2020", "Park et al. 2014",
  "COPD summary", "", "Pletcher et al. 2019", "Soret et al. 2020",
  "Narayanamurthy et al. 2017", "Coburn et al. 2015", "Carmody et al. 2013",
  "", "CF summary", "", "Byun et al. 2017", "", "Summary"))

coef <- with(test_shannon, cbind(coef1, coef3, coef4, coef2))
low <- with(test_shannon, cbind(low1, low3, low4, low2))
high <- with(test_shannon, cbind(high1, high3, high4, high2))
pdf(file = "difference_shannon_BC.pdf", width = 12, height = 10)
forestplot(row_names, coef, low, high, hrzl_lines = list(`2` = gpar(lty = 1),

```

```

`16` = gpar(lty = 2), `27` = gpar(lty = 2), `35` = gpar(lty = 2),
`37` = gpar(lty = 1)), title = "Difference of mean Shannon 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, 5), TRUE, rep(FALSE, 2), TRUE, TRUE, rep(FALSE,
  3), TRUE, rep(FALSE, 6), TRUE, rep(FALSE, 6), TRUE,
  TRUE, rep(FALSE, 3), TRUE), xlab = "Difference of mean Shannon index",
new_page = FALSE, legend = c("Healthy/Stable", "Healthy/Diseased",
  "Stable/Exacerbated", "Healthy/Exacerbated"), legend_args = fpLegend(pos =
  ↪ list(x = 0.85,
  y = 0.85), title = "Difference", r = unit(0.1, "snpc"),
  gp = gpar(col = "#CCCCCC", lwd = 1.5)))
dev.off()

```