

```

# Digital CAP validation Routine HE slides
# Author: Anand Rajan
# Started: 3-18-2020

#agreea function - Agree function
#how to use: agreea(data.frame[,columns])
agreea<-function (ratings, tolerance = 0){
  ratings <- as.matrix(na.omit(ratings))
  ns <- nrow(ratings)
  nr <- ncol(ratings)
  if (is.numeric(ratings)) {
    rangetab <- apply(ratings, 1, max) - apply(ratings, 1, min)
    coeff <- 100 * sum(rangetab <= tolerance)/ns
  }
  else {
    rangetab <- as.numeric(sapply(apply(ratings, 1, table), length))
    coeff <- 100 * (sum(rangetab == 1)/ns)
    tolerance <- 0
  }
  return(coeff)}
#install ggthemr using:
#devtools::install_github('cttobin/ggthemr')
libs<-c('ggplot2', 'ggthemr', 'readxl', 'Rmisc')
lapply(libs, require, character.only = TRUE)

#### Full case set analysis

# clean it out; initialize variables
agreements<-NULL
agreementsA<-NULL
u<-NULL
v<-NULL
dresultP<-NULL

```

```

dresultPMaj<-NULL
# vector to pull 90 random numbers from
a<-1:171
# vector of repeats standing for gold standard glass diagnosis
dglass<-rep(1, 171)
# just an index
index<-1:171
agreements<-vector(length=1000) #result-collecting vector
agreementsA<-vector(length=1000) #result-collecting vector
#build diagnosis table for pathologist.
#dresultP = pathologist diagnosis on digital
ALL_Data<-read_excel("Bob Steph Sarag Bellizzi Samuelson Cases.xlsx", sheet =
"All") #read in data
alldata<-as.data.frame(ALL_Data) #change to data.frame
dresultP<-alldata$DigitalBase #digital result Pathologist gets base
concordance
dresultPMaj<-alldata$DigitalMajor #digital result Pathologist Major gets
minor excluded concordance
diagnosisP<-data.frame(cbind(index, dglass, dresultP))
diagnosisPMaj<-data.frame(cbind(index, dglass, dresultPMaj))
#doing it 1000 times
#build 1 x 1 table for random 90 cases using dt[dt$fct %in% vc,]
#calculate percent agreement
#dump and collect percent agreement in vector
#loop
for (i in 1:1000){
  key<-sample(a, 90)
  dsample<-diagnosisP[diagnosisP$index %in% key,]
  agreements[i]<-agreea(dsample[,2:3])}
for (i in 1:1000){
  key<-sample(a, 90)
  dsampleMaj<-diagnosisPMaj[diagnosisPMaj$index %in% key,]
  agreementsA[i]<-agreea(dsampleMaj[,2:3])}
u<-data.frame(agreements)
v<-data.frame(agreementsA)

```

```

pMinMaj<-ggplot(data=u, aes(agreements))+geom_bar() + scale_x_binned()+
  labs(x="Intraobserver agreement (%) with original diagnosis", y = "Count",
       subtitle = "All 5 study pathologists included.",
       caption = "") +
  ggtitle("Distribution of percent agreement (%) in 90 cases (1000 samples).
\nMajor and minor discrepancies counted.") +
  theme(plot.title = element_text(lineheight=.9, face="bold"))
pMaj<-ggplot(data=v, aes(agreementsA))+geom_bar() + scale_x_binned()+
  labs(x="Percent agreement (%) with original diagnosis", y = "Count",
       subtitle = "All 5 study pathologists included.",
       caption = "") +
  ggtitle("Distribution of percent agreement (%) in 90 cases (1000 samples).
\nOnly major discrepancies counted.") +
  theme(plot.title = element_text(lineheight=.9, face="bold"))
#####
#####
#Output Section
#####
#####
#CIs
CI(agreements)
CI(agreementsA)
#graphs
ggthemr('greyscale')
darken_swatch(amount=0.1)
ggsave(pMaj, file ="pMaj.pdf", width=8, height=9.04, dpi=300)
ggsave(pMinMaj, file ="pMinMaj.pdf", width=8, height=9.04, dpi=300)
ggsave(pMaj, file ="pMaj.eps", width=8, height=9.04, dpi=300)
ggsave(pMinMaj, file ="pMinMaj.eps", width=8, height=9.04, dpi=300)

```