Checking Error Caused by Using Masked Carrier Risk

```
library(tidyverse)
load("RObjects/summary_tables/summaryTable10000Families.Rdata")
```

Using All Families

```
# calculate average predicted value
avg_full <- mean(summaryTable$fullCarrierRisk) #adjusted means adjusted via prediction model
avg_masked <- mean(summaryTable$carrierRiskUnaffectedInfoMasked)
avg_true <- mean(as.numeric(summaryTable$probandMLH1Status))

sum_full <- sum(summaryTable$fullCarrierRisk) #adjusted means adjusted via prediction model
sum_masked <- sum(summaryTable$carrierRiskUnaffectedInfoMasked)
sum_true <- sum(as.numeric(summaryTable$probandMLH1Status))
# print average predicted value
print(str_interp("Average full carrier risk score: ${avg_full}"))
## [1] "Average full carrier risk score: 0.08222868608331"
print(str_interp("Average masked (unadjusted) carrier risk score: ${avg_masked}"))
## [1] "Average masked (unadjusted) carrier risk score: 0.138178890715236"
print(str_interp("Average number of true carriers of MLH1: ${avg_true}"))
## [1] "Average number of true carriers of MLH1: 0.0978"

mse_full <- mean((as.numeric(summaryTable$probandMLH1Status) - summaryTable$fullCarrierRisk)**2)
mse_masked <- mean((as.numeric(summaryTable$probandMLH1Status) - summaryTable$carrierRiskUnaffectedInfo
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.0594202717519354"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.0612137433039341"
```

Using Families with Unaffected First Degree Relatives (U1)

Unrounded

```
summaryTableFirstDegUnaff <- summaryTable %>% filter(firstDegreeAffectedFamilyMembersBinary == 0)


# calculate average predicted value
avg_full <- mean(summaryTableFirstDegUnaff$fullCarrierRisk) #adjusted means adjusted via prediction mod
avg_masked <- mean(summaryTableFirstDegUnaff$carrierRiskUnaffectedInfoMasked)
avg_true <- mean(as.numeric(summaryTableFirstDegUnaff$probandMLH1Status))
# print average predicted value
print(str_interp("Average full carrier risk score: ${avg_full}"))
## [1] "Average full carrier risk score: 0.0332079222774842"
print(str_interp("Average masked  carrier risk score: ${avg_masked}"))
## [1] "Average masked  carrier risk score: 0.095801099097885"
```

```
print(str_interp("Average number of true carriers of MLH1: ${avg_true}"))
## [1] "Average number of true carriers of MLH1: 0.046973803071364"

mse_full <- mean((as.numeric(summaryTableFirstDegUnaff$probandMLH1Status) - summaryTableFirstDegUnaff$fu
mse_masked <- mean((as.numeric(summaryTableFirstDegUnaff$probandMLH1Status) - summaryTableFirstDegUnaff$
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.0369271312819115"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.0389315218152308"
```

Using Families with Affected First Degree Relatives (A1)

```
summaryTableFirstDegAff <- summaryTable %>% filter(firstDegreeAffectedFamilyMembersBinary == 1)


# calculate average predicted value
avg_full <- mean(summaryTableFirstDegAff$fullCarrierRisk) #adjusted means adjusted via prediction model
avg_masked <- mean(summaryTableFirstDegAff$carrierRiskUnaffectedInfoMasked)
avg_true <- mean(as.numeric(summaryTableFirstDegAff$probandMLH1Status))
# print average predicted value
print(str_interp("Average full carrier risk score: ${avg_full}"))
## [1] "Average full carrier risk score: 0.461711102398339"
print(str_interp("Average masked  carrier risk score: ${avg_masked}"))
## [1] "Average masked  carrier risk score: 0.47026945261475"
print(str_interp("Average number of true carriers of MLH1: ${avg_true}"))
## [1] "Average number of true carriers of MLH1: 0.491258741258741"

mse_full <- mean((as.numeric(summaryTableFirstDegAff$probandMLH1Status) - summaryTableFirstDegAff$fullCa
mse_masked <- mean((as.numeric(summaryTableFirstDegAff$probandMLH1Status) - summaryTableFirstDegAff$cari
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.233545492033869"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.233706185178021"
```

Round values with clinical significance

```
round <- function(x){
  if(x>=0.2){
    return(1)
  }
  else{
    return(0)
  }
}

summaryTable$roundedFullCarrierRisk <- lapply(summaryTable$fullCarrierRisk, round)
summaryTable$roundedFullCarrierRisk <- as.numeric(summaryTable$roundedFullCarrierRisk)
summaryTable$roundedCarrierRiskUnaffectedInfoMasked <- lapply(summaryTable$carrierRiskUnaffectedInfoMask
summaryTable$roundedCarrierRiskUnaffectedInfoMasked <- as.numeric(summaryTable$roundedCarrierRiskUnaffec
summaryTable$probandMLH1Status <- as.numeric(summaryTable$probandMLH1Status)
```

```r
#All families

#Calibration
print(str_interp("Sum of rounded full carrier risk: ${sum(summaryTable$roundedFullCarrierRisk)}"))
## [1] "Sum of rounded full carrier risk: 1167"
print(str_interp("Sum of rounded masked carrier risk: ${sum(summaryTable$roundedCarrierRiskUnaffectedIn
## [1] "Sum of rounded masked carrier risk: 2340"
print(str_interp("Sum of true MLH1 carriers: ${sum(summaryTable$probandMLH1Status)}"))
## [1] "Sum of true MLH1 carriers: 978"

mse_full <- mean((as.numeric(summaryTable$probandMLH1Status) - summaryTable$roundedFullCarrierRisk)**2)
mse_masked <- mean((as.numeric(summaryTable$probandMLH1Status) - summaryTable$roundedCarrierRiskUnaffec
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.0947"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.1716"

#First Degree Unaffected
summaryTableFirstDegUnaff <- summaryTable %>% filter(firstDegreeAffectedFamilyMembersBinary == 0)

#Calibration
print(str_interp("Sum of rounded full carrier risk: ${sum(summaryTableFirstDegUnaff$roundedFullCarrierR
## [1] "Sum of rounded full carrier risk: 79"
print(str_interp("Sum of rounded masked carrier risk: ${sum(summaryTableFirstDegUnaff$roundedCarrierRisk
## [1] "Sum of rounded masked carrier risk: 1223"
print(str_interp("Sum of true MLH1 carriers: ${sum(summaryTableFirstDegUnaff$probandMLH1Status)}"))
## [1] "Sum of true MLH1 carriers: 416"

mse_full <- mean((as.numeric(summaryTableFirstDegUnaff$probandMLH1Status) - summaryTableFirstDegUnaff$ro
mse_masked <- mean((as.numeric(summaryTableFirstDegUnaff$probandMLH1Status) - summaryTableFirstDegUnaff$
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.0446025293586269"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.129516711833785"

#First degree affected
summaryTableFirstDegAff <- summaryTable %>% filter(firstDegreeAffectedFamilyMembersBinary == 1)
#Calibration
print(str_interp("Sum of rounded full carrier risk: ${sum(summaryTableFirstDegAff$roundedFullCarrierRisk
## [1] "Sum of rounded full carrier risk: 1088"
print(str_interp("Sum of rounded masked carrier risk: ${sum(summaryTableFirstDegAff$roundedCarrierRiskUn
## [1] "Sum of rounded masked carrier risk: 1117"
print(str_interp("Sum of true MLH1 carriers: ${sum(summaryTableFirstDegAff$probandMLH1Status)}"))
## [1] "Sum of true MLH1 carriers: 562"

mse_full <- mean((as.numeric(summaryTableFirstDegAff$probandMLH1Status) - summaryTableFirstDegAff$rounde
mse_masked <- mean((as.numeric(summaryTableFirstDegAff$probandMLH1Status) - summaryTableFirstDegAff$roun
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.482517482517482"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.497377622377622"
```