

Checking Error Caused by Using Masked Carrier Risk

```
library(tidyverse)
load("RObjects/summary_tables/summaryTable10000Families_no_censoring.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: 0.0968539608410206"
print(str_interp("Average masked (unadjusted) carrier risk score: ${avg_masked}"))
## [1] "Average masked (unadjusted) carrier risk score: 0.244555038563386"
print(str_interp("Average number of true carriers of MLH1: ${avg_true}"))
## [1] "Average number of true carriers of MLH1: 0.0987"

mse_full <- mean((as.numeric(summaryTable$probandMLH1Status) - summaryTable$fullCarrierRisk)**2)
mse_masked <- mean((as.numeric(summaryTable$probandMLH1Status) - summaryTable$carrierRiskUnaffectedInfoMasked)**2)
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.0143830140600578"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.057399747853149"
```

Using Families with Unaffected First Degree Relatives (U1)

Unrounded

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

# calculate average predicted value
avg_full <- mean(summaryTableFirstDegUnaff$fullCarrierRisk) #adjusted means adjusted via prediction model
avg_masked <- mean(summaryTableFirstDegUnaff$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: 0.00342436866470362"
print(str_interp("Average masked carrier risk score: ${avg_masked}"))
## [1] "Average masked carrier risk score: 0.138875464251585"
```

```

print(str_interp("Average number of true carriers of MLH1: ${avg_true}"))
## [1] "Average number of true carriers of MLH1: 0.00305343511450382"

mse_full <- mean((as.numeric(summaryTableFirstDegUnaff$probandMLH1Status) - summaryTableFirstDegUnaff$fullCarrierRisk))
mse_masked <- mean((as.numeric(summaryTableFirstDegUnaff$probandMLH1Status) - summaryTableFirstDegUnaff$carrierRiskUnaffectedInfoMasked))
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.000750616302085828"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.0379542609721075"

```

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: 0.231023862395074"
print(str_interp("Average masked carrier risk score: ${avg_masked}"))
## [1] "Average masked carrier risk score: 0.39631657097948"
print(str_interp("Average number of true carriers of MLH1: ${avg_true}"))
## [1] "Average number of true carriers of MLH1: 0.23605359317905"

mse_full <- mean((as.numeric(summaryTableFirstDegAff$probandMLH1Status) - summaryTableFirstDegAff$fullCarrierRisk))
mse_masked <- mean((as.numeric(summaryTableFirstDegAff$probandMLH1Status) - summaryTableFirstDegAff$carrierRiskUnaffectedInfoMasked))
print(str_interp("Full MSE: ${mse_full}"))
## [1] "Full MSE: 0.0339598678440395"
print(str_interp("Masked MSE: ${mse_masked}"))
## [1] "Masked MSE: 0.0853245091597847"

```

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$roundedCarrierRiskUnaffectedInfoMasked <- lapply(summaryTable$carrierRiskUnaffectedInfoMasked, round)
summaryTable$roundedCarrierRiskUnaffectedInfoMasked <- as.numeric(summaryTable$roundedCarrierRiskUnaffectedInfoMasked)
summaryTable$probandMLH1Status <- as.numeric(summaryTable$probandMLH1Status)

```

```
#All families
```

```
#Calibration
```

```
print(str_interp("Sum of rounded full carrier risk: ${sum(summaryTable$roundedFullCarrierRisk)}"))  
## [1] "Sum of rounded full carrier risk: 1057"  
print(str_interp("Sum of rounded masked carrier risk: ${sum(summaryTable$roundedCarrierRiskUnaffectedIn)}"))  
## [1] "Sum of rounded masked carrier risk: 3964"  
print(str_interp("Sum of true MLH1 carriers: ${sum(summaryTable$probandMLH1Status)}"))  
## [1] "Sum of true MLH1 carriers: 987"
```

```
mse_full <- mean((as.numeric(summaryTable$probandMLH1Status) - summaryTable$roundedFullCarrierRisk)**2)  
mse_masked <- mean((as.numeric(summaryTable$probandMLH1Status) - summaryTable$roundedCarrierRiskUnaffectedIn)**2)  
print(str_interp("Full MSE: ${mse_full}"))  
## [1] "Full MSE: 0.0216"  
print(str_interp("Masked MSE: ${mse_masked}"))  
## [1] "Masked MSE: 0.2995"
```

```
#First Degree Unaffected
```

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

```
#Calibration
```

```
print(str_interp("Sum of rounded full carrier risk: ${sum(summaryTableFirstDegUnaff$roundedFullCarrierRisk)}"))  
## [1] "Sum of rounded full carrier risk: 20"  
print(str_interp("Sum of rounded masked carrier risk: ${sum(summaryTableFirstDegUnaff$roundedCarrierRiskUnaffectedIn)}"))  
## [1] "Sum of rounded masked carrier risk: 1199"  
print(str_interp("Sum of true MLH1 carriers: ${sum(summaryTableFirstDegUnaff$probandMLH1Status)}"))  
## [1] "Sum of true MLH1 carriers: 18"
```

```
mse_full <- mean((as.numeric(summaryTableFirstDegUnaff$probandMLH1Status) - summaryTableFirstDegUnaff$roundedFullCarrierRisk)**2)  
mse_masked <- mean((as.numeric(summaryTableFirstDegUnaff$probandMLH1Status) - summaryTableFirstDegUnaff$roundedCarrierRiskUnaffectedIn)**2)  
print(str_interp("Full MSE: ${mse_full}"))  
## [1] "Full MSE: 0.00135708227311281"  
print(str_interp("Masked MSE: ${mse_masked}"))  
## [1] "Masked MSE: 0.200339270568278"
```

```
#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: 1037"  
print(str_interp("Sum of rounded masked carrier risk: ${sum(summaryTableFirstDegAff$roundedCarrierRiskUnaffectedIn)}"))  
## [1] "Sum of rounded masked carrier risk: 2765"  
print(str_interp("Sum of true MLH1 carriers: ${sum(summaryTableFirstDegAff$probandMLH1Status)}"))  
## [1] "Sum of true MLH1 carriers: 969"
```

```
mse_full <- mean((as.numeric(summaryTableFirstDegAff$probandMLH1Status) - summaryTableFirstDegAff$roundedFullCarrierRisk)**2)  
mse_masked <- mean((as.numeric(summaryTableFirstDegAff$probandMLH1Status) - summaryTableFirstDegAff$roundedCarrierRiskUnaffectedIn)**2)  
print(str_interp("Full MSE: ${mse_full}"))  
## [1] "Full MSE: 0.0506699147381242"  
print(str_interp("Masked MSE: ${mse_masked}"))  
## [1] "Masked MSE: 0.44190012180268"
```