Part1 Descriptive Means Tab;e

Kim Tiller

April 19, 2021

Totals - Prevalence and Cost

```r
totals <- read.csv("Customer_Totals.csv")
#Commercial
comm <-
  totals %>% filter(plan_type=="Commercial") %>%
  group_by(db,year) %>%
  summarise(mbrs=sum(total_members),
            cost=sum(diab_medical+nd_medical+diab_pharmacy+nd_pharmacy),
            diab_mbrs=sum(diab_members),
            diab_cost=sum(diab_medical+diab_pharmacy),
            .groups = "keep") %>%
  mutate(diab_prev = diab_mbrs/mbrs,
                diab_pctcost = diab_cost/cost)
#Mcaid
mcaid <-
  totals %>% filter(plan_type=="Medicaid") %>%
  group_by(db,year) %>%
  summarise(mbrs=sum(total_members),
            cost=sum(diab_medical+nd_medical+diab_pharmacy+nd_pharmacy),
            diab_mbrs=sum(diab_members),
            diab_cost=sum(diab_medical+diab_pharmacy),
            .groups = "keep") %>%
  mutate(diab_prev = diab_mbrs/mbrs,
                diab_pctcost = diab_cost/cost)
#MCare
mcare <-
  totals %>% filter(plan_type=="Medicare") %>%
  group_by(db,year) %>%
  summarise(mbrs=sum(total_members),
            cost=sum(diab_medical+nd_medical+diab_pharmacy+nd_pharmacy),
            diab_mbrs=sum(diab_members),
            diab_cost=sum(diab_medical+diab_pharmacy),
            .groups = "keep") %>%
  mutate(diab_prev = diab_mbrs/mbrs,
                diab_pctcost = diab_cost/cost)
```

Utilization

```r
usage <- read.csv("Customer_utilization.csv")
usage_join <- inner_join(usage,totals,by=c("db" = "db", "year" = "year", "plan_type" = "plan_type", "ag

usage_calc <-
  usage_join %>% group_by(db,year,plan_type) %>%
  summarise(diab_mbrs=sum(diab_members),
            diab_er=sum(diab_er_visits),
            diab_ip=sum(diab_ip_days),
            diab_op=sum(diab_op_visits),
            diab_cost=sum(diab_medical+diab_pharmacy),
            nd_mbrs=sum(nd_members),
            nd_er=sum(nd_er_visits),
            nd_ip=sum(nd_ip_days),
            nd_op=sum(nd_op_visits),
            nd_cost=sum(nd_medical+nd_pharmacy), .groups="keep") %>%
  mutate(diab_er1000 = (diab_er/diab_mbrs)*1000,
         diab_ip1000 = (diab_ip/diab_mbrs)*1000,
         diab_oppm = (diab_op/diab_mbrs),
         diab_pmpy = diab_cost/diab_mbrs,
         nd_er1000 = (nd_er/nd_mbrs)*1000,
         nd_ip1000 = (nd_ip/nd_mbrs)*1000,
         nd_oppm = (nd_op/nd_mbrs),
         nd_pmpy = nd_cost/nd_mbrs)
```

Risk

```r
risk <- read.csv("Customer_risk.csv")
risk_calc <- risk %>%
  group_by(db,year,plan_type) %>%
  summarise(diab_mbr=sum(diab_mbr),
            diab_risk=sum(diab_risk_score),
            nd_mbr=sum(nd_mbr),
            nd_risk=sum(nd_risk_score),
            .groups = "keep")%>%
            mutate(diab_risk_avg = diab_risk/diab_mbr,
                   nd_risk_avg = nd_risk/nd_mbr)
```

Quality

```r
quality <- read.csv("Customer_quality.csv")
quality_calc <- quality %>%
  group_by(db,year,plan_type) %>%
  summarise(HbaTest_denom = sum(HbaTest_denom),
            HbaTest_numer = sum(HbaTest_numer),
            Hba1c8_denom = sum(Hba1c8_denom),
            Hba1c8_numer = sum(Hba1c8_numer),
```

```r
            eye_denom = sum(eye_denom),
            eye_numer = sum(eye_numer),
            neph_denom = sum(neph_denom),
            neph_numer = sum(neph_numer),
            statT_denom = sum(statT_denom),
            statT_numer = sum(statT_numer),
            statA_denom = sum(statA_denom),
            statA_numer = sum(statA_numer),
            .groups = "keep") %>%
            mutate(HbaTest = HbaTest_numer/HbaTest_denom,
                   Hba1c8 = Hba1c8_numer/Hba1c8_denom,
                   eye_exam = eye_numer/eye_denom,
                   neph = neph_numer/neph_denom,
                   statT = statT_numer/statT_denom,
                   statA = statA_numer/statA_denom)
```

Mean and CI

```r
#Prevalence%
c_prev <- t.test(comm$diab_prev, conf.level = .95)
md_prev <- t.test(mcaid$diab_prev, conf.level = .95)
mr_prev <- t.test(mcare$diab_prev, conf.level = .95)
#Cost %
c_cost <- t.test(comm$diab_pctcost, conf.level = .95)
md_cost <- t.test(mcaid$diab_pctcost, conf.level = .95)
mr_cost <- t.test(mcare$diab_pctcost, conf.level = .95)


#Diab ER
c_diab_er <- t.test(usage_calc$diab_er1000[usage_calc$plan_type=='Commercial'], conf.level = .95)
md_diab_er <- t.test(usage_calc$diab_er1000[usage_calc$plan_type=='Medicaid'], conf.level = .95)
mr_diab_er <- t.test(usage_calc$diab_er1000[usage_calc$plan_type=='Medicare'], conf.level = .95)
#Diab IP
c_diab_ip <- t.test(usage_calc$diab_ip1000[usage_calc$plan_type=='Commercial'], conf.level = .95)
md_diab_ip <- t.test(usage_calc$diab_ip1000[usage_calc$plan_type=='Medicaid'], conf.level = .95)
mr_diab_ip <- t.test(usage_calc$diab_ip1000[usage_calc$plan_type=='Medicare'], conf.level = .95)
#Diab OP
c_diab_op <- t.test(usage_calc$diab_oppm[usage_calc$plan_type=='Commercial'], conf.level = .95)
md_diab_op <- t.test(usage_calc$diab_oppm[usage_calc$plan_type=='Medicaid'], conf.level = .95)
mr_diab_op <- t.test(usage_calc$diab_oppm[usage_calc$plan_type=='Medicare'], conf.level = .95)
#Diab PMPY
c_diab_pmpy <- t.test(usage_calc$diab_pmpy[usage_calc$plan_type=='Commercial'], conf.level = .95)
md_diab_pmpy <- t.test(usage_calc$diab_pmpy[usage_calc$plan_type=='Medicaid'], conf.level = .95)
mr_diab_pmpy <- t.test(usage_calc$diab_pmpy[usage_calc$plan_type=='Medicare'], conf.level = .95)


#NONDiab ER
c_ND_er <- t.test(usage_calc$nd_er1000[usage_calc$plan_type=='Commercial'], conf.level = .95)
md_ND_er <- t.test(usage_calc$nd_er1000[usage_calc$plan_type=='Medicaid'], conf.level = .95)
mr_ND_er <- t.test(usage_calc$nd_er1000[usage_calc$plan_type=='Medicare'], conf.level = .95)
#NONDiab IP
c_ND_ip <- t.test(usage_calc$nd_ip1000[usage_calc$plan_type=='Commercial'], conf.level = .95)
md_ND_ip <- t.test(usage_calc$nd_ip1000[usage_calc$plan_type=='Medicaid'], conf.level = .95)
```

```r
mr_ND_ip <- t.test(usage_calc$nd_ip1000[usage_calc$plan_type=='Medicare'], conf.level = .95)
#NONDiab OP
c_ND_op <- t.test(usage_calc$nd_oppm[usage_calc$plan_type=='Commercial'], conf.level = .95)
md_ND_op <- t.test(usage_calc$nd_oppm[usage_calc$plan_type=='Medicaid'], conf.level = .95)
mr_ND_op <- t.test(usage_calc$nd_oppm[usage_calc$plan_type=='Medicare'], conf.level = .95)
#NONDiab PMPY
c_ND_pmpy <- t.test(usage_calc$nd_pmpy[usage_calc$plan_type=='Commercial'], conf.level = .95)
md_ND_pmpy <- t.test(usage_calc$nd_pmpy[usage_calc$plan_type=='Medicaid'], conf.level = .95)
mr_ND_pmpy <- t.test(usage_calc$nd_pmpy[usage_calc$plan_type=='Medicare'], conf.level = .95)


#Diab Risk
c_diab_risk <- t.test(risk_calc$diab_risk_avg[risk_calc$plan_type=='Commercial'], conf.level = .95)
md_diab_risk <- t.test(risk_calc$diab_risk_avg[risk_calc$plan_type=='Medicaid'], conf.level = .95)
mr_diab_risk <- t.test(risk_calc$diab_risk_avg[risk_calc$plan_type=='Medicare'], conf.level = .95)
#NONDiab Risk
c_ND_risk <- t.test(risk_calc$nd_risk_avg[risk_calc$plan_type=='Commercial'], conf.level = .95)
md_ND_risk <- t.test(risk_calc$nd_risk_avg[risk_calc$plan_type=='Medicaid'], conf.level = .95)
mr_ND_risk <- t.test(risk_calc$nd_risk_avg[risk_calc$plan_type=='Medicare'], conf.level = .95)


#Quaility
c_hbatest <- t.test(quality_calc$HbaTest[quality_calc$plan_type=='Commercial'], conf.level = .95)
md_hbatest <- t.test(quality_calc$HbaTest[quality_calc$plan_type=='Medicaid'], conf.level = .95)
mr_hbatest <- t.test(quality_calc$HbaTest[quality_calc$plan_type=='Medicare'], conf.level = .95)


c_hba1c8 <- t.test(quality_calc$Hba1c8[quality_calc$plan_type=='Commercial'], conf.level = .95)
md_hba1c8 <- t.test(quality_calc$Hba1c8[quality_calc$plan_type=='Medicaid'], conf.level = .95)
mr_hba1c8 <- t.test(quality_calc$Hba1c8[quality_calc$plan_type=='Medicare'], conf.level = .95)


c_eye <- t.test(quality_calc$eye_exam[quality_calc$plan_type=='Commercial'], conf.level = .95)
md_eye <- t.test(quality_calc$eye_exam[quality_calc$plan_type=='Medicaid'], conf.level = .95)
mr_eye <- t.test(quality_calc$eye_exam[quality_calc$plan_type=='Medicare'], conf.level = .95)


c_neph <- t.test(quality_calc$neph[quality_calc$plan_type=='Commercial'], conf.level = .95)
md_neph <- t.test(quality_calc$neph[quality_calc$plan_type=='Medicaid'], conf.level = .95)
mr_neph <- t.test(quality_calc$neph[quality_calc$plan_type=='Medicare'], conf.level = .95)


c_statT <- t.test(quality_calc$statT[quality_calc$plan_type=='Commercial'], conf.level = .95)
md_statT <- t.test(quality_calc$statT[quality_calc$plan_type=='Medicaid'], conf.level = .95)
mr_statT <- t.test(quality_calc$statT[quality_calc$plan_type=='Medicare'], conf.level = .95)


c_statA <- t.test(quality_calc$statA[quality_calc$plan_type=='Commercial'], conf.level = .95)
md_statA <- t.test(quality_calc$statA[quality_calc$plan_type=='Medicaid'], conf.level = .95)
mr_statA <- t.test(quality_calc$statA[quality_calc$plan_type=='Medicare'], conf.level = .95)


#Commercial Data
comm_desc <- c("comm prev", "comm cost", "comm diab er", "comm diab ip", "comm diab op", "comm diab pmpy
          ,"comm ND er","comm ND ip", "comm ND op", "comm ND pmpy", "comm diab risk", "comm ND risk"
          ,"comm Hba1Test", "comm Hba1c8", "comm eye", "comm neph", "comm statT", "comm statA")
comm_table <- map_df(list(c_prev, c_cost, c_diab_er, c_diab_ip, c_diab_op, c_diab_pmpy
                     ,c_ND_er ,c_ND_ip, c_ND_op, c_ND_pmpy, c_diab_risk, c_ND_risk
                     ,c_hbatest, c_hba1c8, c_eye, c_neph, c_statT, c_statA), tidy)
comm_final <- cbind(comm_desc, comm_table[c("estimate", "conf.low", "conf.high", "p.value")])
```

```r
#Medicaid Data
mcaid_desc <- c("mcaid prev", "mcaid cost", "mcaid diab er", "mcaid diab ip", "mcaid diab op", "mcaid di
                , "mcaid ND er", "mcaid ND ip", "mcaid ND op", "mcaid ND pmpy", "mcaid diab risk", "mcai
                ,"mcaid Hba1Test", "mcaid Hba1c8", "mcaid eye", "mcaid neph", "mcaid statT", "mcaid stat
mcaid_table <- map_df(list(md_prev, md_cost, md_diab_er, md_diab_ip, md_diab_op, md_diab_pmpy
                      , md_ND_er, md_ND_ip, md_ND_op, md_ND_pmpy, md_diab_risk, md_ND_risk
                      ,md_hbatest, md_hba1c8, md_eye, md_neph, md_statT, md_statA), tidy)
mcaid_final <- cbind(mcaid_desc, mcaid_table[c("estimate", "conf.low", "conf.high", "p.value")])

#Medicare Data
mcare_desc <- c("mcare prev", "mcare cost", "mcare diab er", "mcare diab ip", "mcare diab op", "mcare di
                , "mcare ND er", "mcare ND ip", "mcare ND op", "mcare ND pmpy", "mcare diab risk", "mcai
                ,"mcare Hba1Test", "mcare Hba1c8", "mcare eye", "mcare neph", "mcare statT", "mcare stat
mcare_table <- map_df(list(mr_prev, mr_cost, mr_diab_er, mr_diab_ip, mr_diab_op, mr_diab_pmpy
                      , mr_ND_er, mr_ND_ip, mr_ND_op, mr_ND_pmpy, mr_diab_risk, mr_ND_risk
                      ,mr_hbatest, mr_hba1c8, mr_eye, mr_neph, mr_statT, mr_statA), tidy)
mcare_final <- cbind(mcare_desc, mcare_table[c("estimate", "conf.low", "conf.high", "p.value")])
```

Write to files

```r
write.csv(comm_final, "table_comm.csv")
write.csv(mcaid_final, "table_mcaid.csv")
write.csv(mcare_final, "table_mcare.csv")
```

Testing for Normal Distribution

```r
shapiro.test(comm$diab_prev)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  comm$diab_prev
## W = 0.90613, p-value = 0.001137
```

```r
shapiro.test(mcaid$diab_prev)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  mcaid$diab_prev
## W = 0.95996, p-value = 0.6009
```

```r
shapiro.test(mcare$diab_prev)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  mcare$diab_prev
## W = 0.93375, p-value = 0.0618
```

```
shapiro.test(comm$diab_pctcost)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  comm$diab_pctcost
## W = 0.97842, p-value = 0.5286
```

```
shapiro.test(mcaid$diab_pctcost)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  mcaid$diab_pctcost
## W = 0.91501, p-value = 0.1055
```

```
shapiro.test(mcare$diab_pctcost)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  mcare$diab_pctcost
## W = 0.96815, p-value = 0.4898
```

```
shapiro.test(usage_calc$diab_er1000[usage_calc$plan_type=='Commercial'])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  usage_calc$diab_er1000[usage_calc$plan_type == "Commercial"]
## W = 0.94171, p-value = 0.09202
```

```
shapiro.test(usage_calc$diab_er1000[usage_calc$plan_type=='Medicaid'])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  usage_calc$diab_er1000[usage_calc$plan_type == "Medicaid"]
## W = 0.88244, p-value = 0.09419
```

```
shapiro.test(usage_calc$diab_er1000[usage_calc$plan_type=='Medicare'])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  usage_calc$diab_er1000[usage_calc$plan_type == "Medicare"]
## W = 0.94684, p-value = 0.3217
```

Commercial data do not pass the test for normal distribution. I believe this is because one commercial plan more closely matches medicare populations due to the age of members. The commercial plan includes many retired members. I considered excluding the plan but the membership is small and many of the members are below the age of 65.

Export rollup for validation and spot checking

```r
write.csv(usage_calc, "usage_calc.csv")
write.csv(quality_calc, "quality_calc.csv")
```