Web-Based Supporting Materials for SIM-12-0064 Resubmission
"Extending the Peters-Belson approach for assessing disparities to right censored time-to-event outcomes"

Computer code and output

Here is R code, using the *survival* package, and output for Peters-Belson comparing incident Dual Incontinence (DI) in Hispanics for the case in which the analysis is not stratified by nursing home. The one-sample log rank test is computed using a call to the `glm` function, which fits a Poisson regression, as suggested in the documentation for the `survexp` function. In that Poisson regression, the log standardized mortality ratio is the intercept, and the test of whether that intercept is zero is identical to the one-sample log-rank test.

Code for plots and analysis:

```
### read in data set containing Whites who live    ###
###  in NHs that also have Hispanics               ###
### remove NH ID and person ID columns             ###
whitesH1 <- read.csv("raceDI_5h.csv")
whitesH <- whitesH1[,c(-1,-2)]

### read in data set containing Hispanics who live  ###
###  in those same NHs                              ###
### remove NH ID and person ID columns              ###
hisp1<- read.csv("raceDI_4.csv")
hisp<- hisp1[,c(-1,-2)]

### fit Cox model to Whites using all covariates in ###
###  the input data set                             ###
coxWhites <- coxph(Surv(DI_DAYS, HAS_DI)~ . , data = whitesH,
               na.action = na.omit)
summary(coxWhites)

### compute Kaplan-Meier of DI for the Hispanics    ###
kmHispanics <- survfit(Surv(DI_DAYS, HAS_DI)~1, data = hisp)

### compute expected DI for the Hispanics           ###
fitHispanics <- survexp(DI_DAYS~1, data = hisp,
                  ratetable = coxWhites, conditional = F,
                  times=kmHispanics$time)

### plot Kaplan-Meier of DI vs. follow-up time      ###
### superimpose expected DI vs. follow-up time      ###
pdf("hispanics.pdf")
plot(kmHispanics, mark.time = F, xlab="Follow-up time (days)",
       ylab="Proportion without Dual Incontinence (DI)")
legend(600,0.95,lty=c(1,1),col=c(1,2),
       c("Observed with 95% CL","Predicted"))
lines(fitHispanics, col = 2)
graphics.off()

### compute one-sample likelihood ratio test        ###
```

```
### bound predicted hazards away from 0 by adding   ###
###  a small constant (0.0002) to all               ###
hazH <- -log(survexp(DI_DAYS~1, data = hisp,
            ratetable = coxWhites, conditional = F, cohort= F))
hazHispanics <- hazH + .00002
lrtHispanics <- glm(HAS_DI ~1 + offset(log(hazHispanics)),
                  family = poisson, data = hisp)
summary(lrtHispanics)
```

Output for Cox model fit using data for Whites:

```
Call:
coxph(formula = Surv(DI_DAYS, HAS_DI) ~ ., data = white, na.action = na.omit)

  n= 16230, number of events= 1782

              coef exp(coef)  se(coef)      z Pr(>|z|)
AGE       0.016876  1.017020  0.003306  5.105 3.31e-07 ***
GENDER   -0.199871  0.818837  0.054248 -3.684 0.000229 ***
ADL_7     0.048695  1.049901  0.003843 12.671  < 2e-16 ***
COMM_2    0.088954  1.093030  0.015736  5.653 1.58e-08 ***
RESTRAIN  0.314821  1.370014  0.077992  4.037 5.42e-05 ***
CHARLSON  0.040339  1.041163  0.016535  2.440 0.014705 *
MED_WKS   0.043807  1.044781  0.019723  2.221 0.026339 *
BOWEL    -0.103370  0.901793  0.062786 -1.646 0.099682 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

         exp(coef) exp(-coef) lower .95 upper .95
AGE         1.0170     0.9833    1.0105    1.0236
GENDER      0.8188     1.2212    0.7362    0.9107
ADL_7       1.0499     0.9525    1.0420    1.0578
COMM_2      1.0930     0.9149    1.0598    1.1273
RESTRAIN    1.3700     0.7299    1.1758    1.5963
CHARLSON    1.0412     0.9605    1.0080    1.0755
MED_WKS     1.0448     0.9571    1.0052    1.0860
BOWEL       0.9018     1.1089    0.7974    1.0199

Concordance= 0.644  (se = 0.008 )
Rsquare= 0.017   (max possible= 0.817 )
Likelihood ratio test= 281.3  on 8 df,   p=0
Wald test            = 292.4  on 8 df,   p=0
Score (logrank) test = 301.3  on 8 df,   p=0
```

One-sample Likelihood Ratio Test Output:

```
Call:
glm(formula = HAS_DI ~ 1 + offset(log(hazH2)), family = poisson,
    data = hisp)

Deviance Residuals:
```

```
    Min       1Q   Median       3Q      Max
-1.9331  -0.2912  -0.0682  -0.0140   3.3616

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.07886    0.11180  -0.705    0.481

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 292.57  on 608  degrees of freedom
Residual deviance: 292.57  on 608  degrees of freedom
AIC: 454.57

Number of Fisher Scoring iterations: 7
```