Base Line Summarization:

For the given project we revisit "The ***HF-ACTION (Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training)***" study conducted on a group of 2030 heart failure patients. The study has been first reported in (O'Connor et al., 2009) and was conducted in a total of 82 sites in USA, Canada and France. The study objective was to assess the effect of adding aerobic exercise training to usual care on the patient's cardiovascular outcomes. The primary endpoint was a composite of all-cause death and all-cause hospitalization. conditions. The baseline characteristics for the number of patients are summarized in Table 1 by Control (Usual Care) group, Treatment (Aerobic Exercise) group and Overall. Quantitative variables are typically summarized by median and inter-quartile range (i.e., the range between the 25th and 75th percentiles), while categorical variables are typically summarized by frequency and percentage. The covariates considered for the study includes the Age of the subject in years, the BMI of the subject, the duration (in minutes) of the baseline cardiopulmonary exercise test, the Gender (Sex) of the Subject, status of etiology for the subject, indicator for history of atrial fibrillation or flutter and nationality of the subject.

Table 1: Baseline Characteristics by Control (Usual Care) group, Treatment (Aerobic Exercise) group and Overall

	Control	**Treatment**	**Overall**
# of Subjects	1070 (50.2%)	1060 (49.8%)	2030 (100%)
Age (Yrs.)	59 (51, 68)	59 (51, 67)	59 (51, 68)
BMI	29.7 (25.9, 35.2)	29.9 (25.8, 35)	29.8 (25.9, 35.1)
Duration of Cardiopulmonary Exercise test (Mins.)	9.7 (7, 12.1)	9.6 (6.9, 12)	9.7 (7, 12)
Sex (male)	789 (73.7%)	742 (70%)	1531 (71.9%)
Sex (female)	281 (26.3%)	318 (30%)	599 (28.1%)
etiology (ischemic: 1)	546 (51%)	550 (51.9%)	1096 (51.5%)
etiology (no-ischemic:2)	524 (49%)	510 (48.1%)	1034 (48.5%)
No history of fibrillation (0)	850 (79.4%)	836 (78.9%)	1686 (79.2%)
History of fibrillation (1)	219 (20.5%)	224 (21.1%)	443 (20.8%)
Nationality USA	946 (88.4%)	934 (88.1%)	1880 (88.3%)
Nationality Canada	86 (8%)	89 (8.4%)	175 (8.2%)
Nationality France	38 (3.6%)	37 (3.5%)	75 (3.5%)
Death rate (per yrs)	0.067	0.062	0.064
Hospitalization rate (per yrs)	0.404	0.401	0.403

It can be observed that the number of subjects in the control group (having usual care) and in the treatment group (having aerobic exercise) is well balanced with approximately 50% of the subjects in both the groups. The Nationality indicator was generated by aggregating the number of subjects in all regions of USA together. The other countries included in the study are Canada and France. The nationality covariate is un-balanced with high number of subjects (~88%) belonging to USA. There is also an unbalance seen in the gender/sex covariate with majority of the subjects being male. Along with this the covariate indicating the history of atrial fibrillation or flutter is also un-balanced with high number of subjects having no history of atrial fibrillation (note that the status indicator for atrial fibrillation is missing for a single subject). A good balance is observed in the ischemic versus non-ischemic etiology status of the subjects. Most subjects in the study are above 51 years of age and have a bit un-healthy BMI of above 30 (the BMI of around 5 subjects is missing). The duration of cardiopulmonary exercise tests for the subjects on average is 9.7 mins within all groups also it is missing for around 20 subjects.

COX-PROPORTIONAL HAZARD MODEL FOR HOSPITALIZATION FREE SURVIVAL

To perform a survival, study the cox proportional hazard model was fit on the data for considering the hazard of the hospitalization free survival by considering only time to first event (events can be hospitalization, death or censoring). However, before fitting the Cox-PH model it was necessary to treat the missing records. In order to impute the missing data, the subjects were grouped based on their gender and then the most appropriate parametric distributions were evaluated for each continuous covariate consisting of missing values. The most appropriate parametric model that describes the continuous covariate for each group (based on gender) was decided using the skewness - kurtosis plot proposed by (Cullen and Frey 1999). To demonstrate an example, consider the Cullen Frey plot for the BMI of males which suggest that the BMI for this group follows a Beta Distribution (Figure 1) with parameters (4.21661, 6.74049) and thus this can be used to generate 5 random samples for imputing the missing values. Similarly, missing values for the duration of cardiopulmonary exercise tests was also determined. The missing value in history of atrial fibrillation or flutter were treated by replacing it with ‘0’ based on the rational that majority of the subjects were not having the history of atrial fibrillation or flutter.

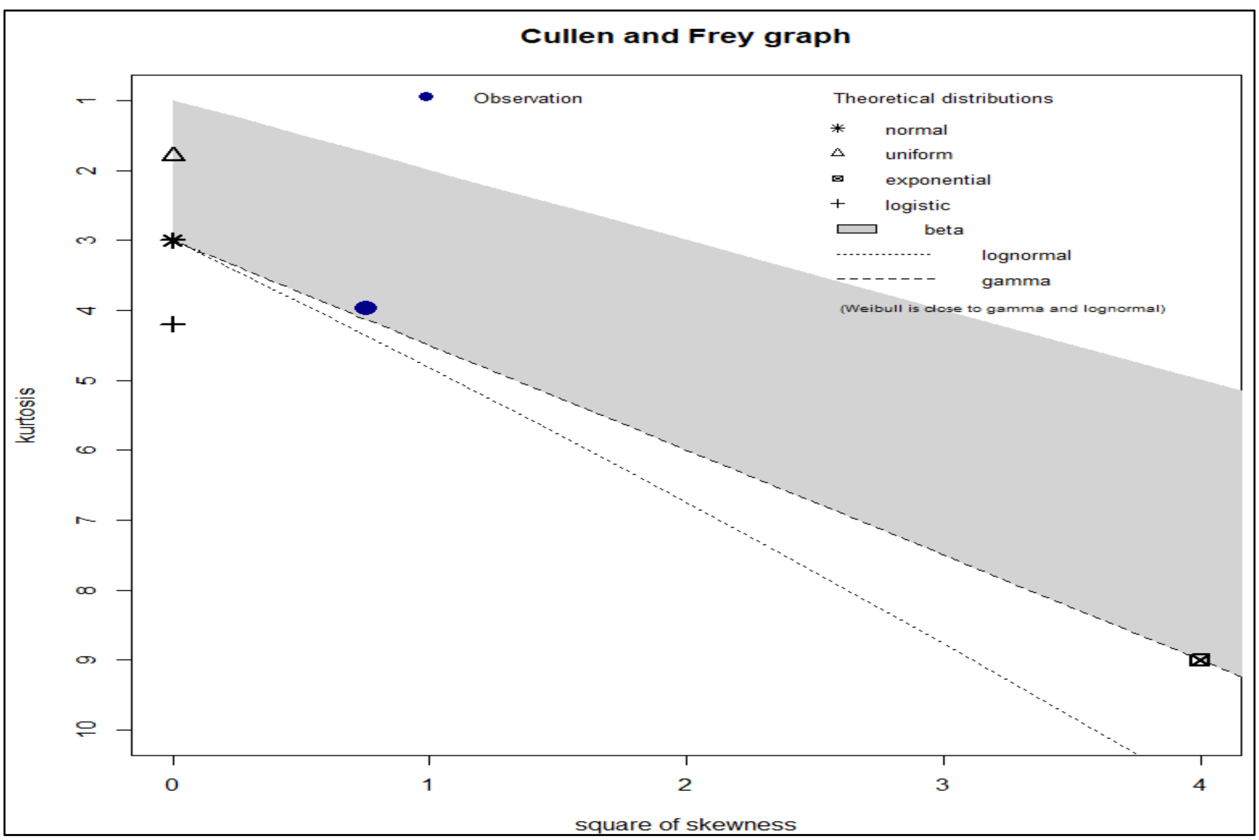

Figure 1: Cullen and Frey plot for deciding the parametric distribution for BMI’s of Male

After having the interpolated data, the Cox-PH model for the time to first event was fit on the data by consider all the 8 covariates (results shown in Figure 2). The age covariate turns out to be non-significant suggesting a need for further investigation about its current functional form. The baseline cumulative hazard function was also evaluated as shown in Figure 3. In order to decide on the goodness of fit for the model the

```
> summary(obj)
Call:
coxph(formula = Surv(time, ttfe_status) ~ factor(trt_ab) + factor(sex) +
    age + factor(etiology) + factor(afibflut) + factor(nationality) +
    bmi + cpxdur, data = df_unique_ss1_interpolated)

  n= 2130, number of events= 1419

                           coef exp(coef)  se(coef)       z Pr(>|z|)
factor(trt_ab)1       -0.109877  0.895945  0.053300  -2.061 0.039257 *
factor(sex)2          -0.219643  0.802805  0.064176  -3.423 0.000620 ***
age                   -0.003726  0.996281  0.002646  -1.408 0.159080
factor(etiology)2     -0.107040  0.898490  0.058465  -1.831 0.067124 .
factor(afibflut)1      0.197813  1.218734  0.065698   3.011 0.002605 **
factor(nationality)2 -0.331838  0.717604  0.101687  -3.263 0.001101 **
factor(nationality)3 -0.708422  0.492421  0.192802  -3.674 0.000238 ***
bmi                   -0.008550  0.991487  0.004136  -2.067 0.038699 *
cpxdur                -0.095918  0.908538  0.008111 -11.826  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                     exp(coef) exp(-coef) lower .95 upper .95
factor(trt_ab)1         0.8959     1.1161    0.8071    0.9946
factor(sex)2            0.8028     1.2456    0.7079    0.9104
age                     0.9963     1.0037    0.9911    1.0015
factor(etiology)2       0.8985     1.1130    0.8012    1.0076
factor(afibflut)1       1.2187     0.8205    1.0715    1.3862
factor(nationality)2    0.7176     1.3935    0.5879    0.8759
factor(nationality)3    0.4924     2.0308    0.3375    0.7185
bmi                     0.9915     1.0086    0.9835    0.9996
cpxdur                  0.9085     1.1007    0.8942    0.9231

Concordance= 0.619  (se = 0.008 )
Likelihood ratio test= 225.8  on 9 df,   p=<2e-16
Wald test            = 217.1  on 9 df,   p=<2e-16
Score (logrank) test = 218.9  on 9 df,   p=<2e-16
```

Figure 2: Cox-PH results for the hospitalization free survival

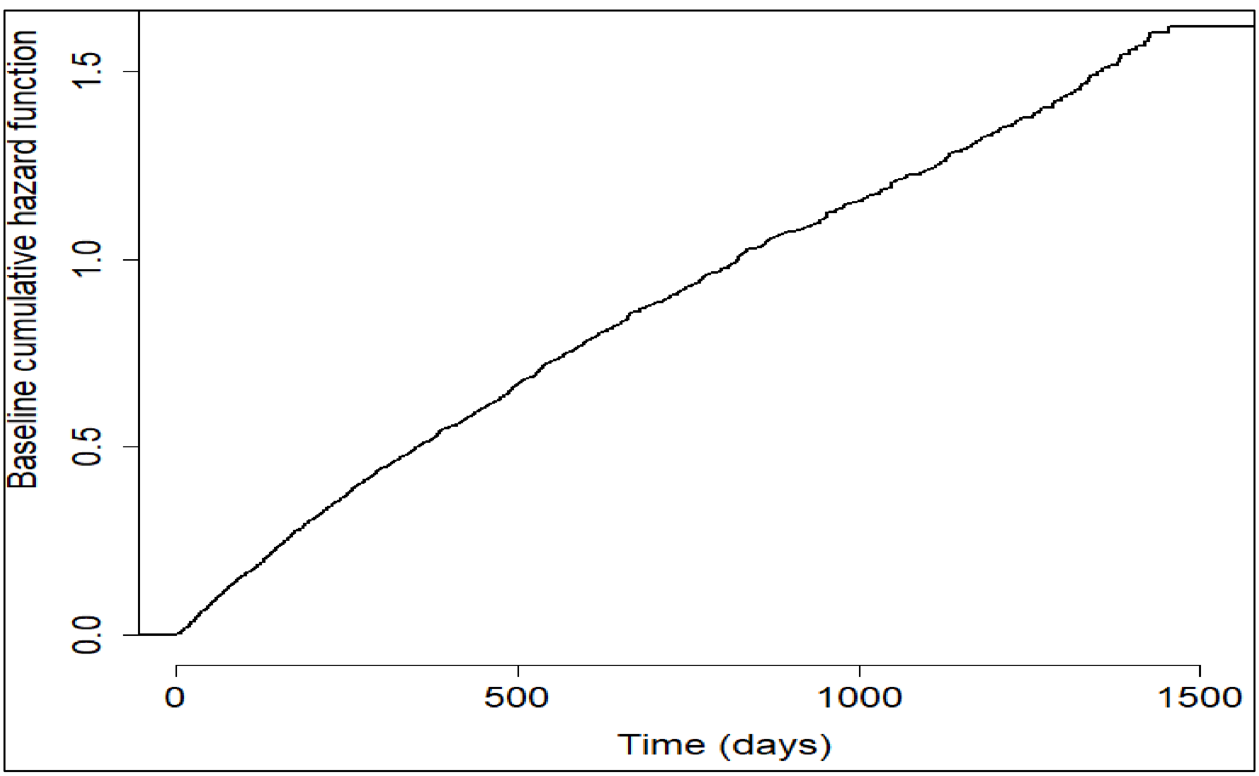

Figure 3: Baseline cumulative hazard for the hospitalization free survival model

In order to decide the over-all goodness of fit for the model the Cox Snell residuals were generated for the model as shown in Figure 4. The straight-line plot suggests that the data fit very well to the model. The Schoenfeld residuals are also generated for identifying if there is any non-proportionality present in each covariate or not. The results of the Schoenfeld residuals are shown in Figure 5 and 6 which suggest that presently the assumption of Non-proportionality is violated by the “Age” covariate significantly along with some significance violation by “BMI” and “Nationality” covariates. Also, the global test for non-proportionality fails for the current functional form of data. Thus, a further investigation into the functional forms of these covariate is necessary.

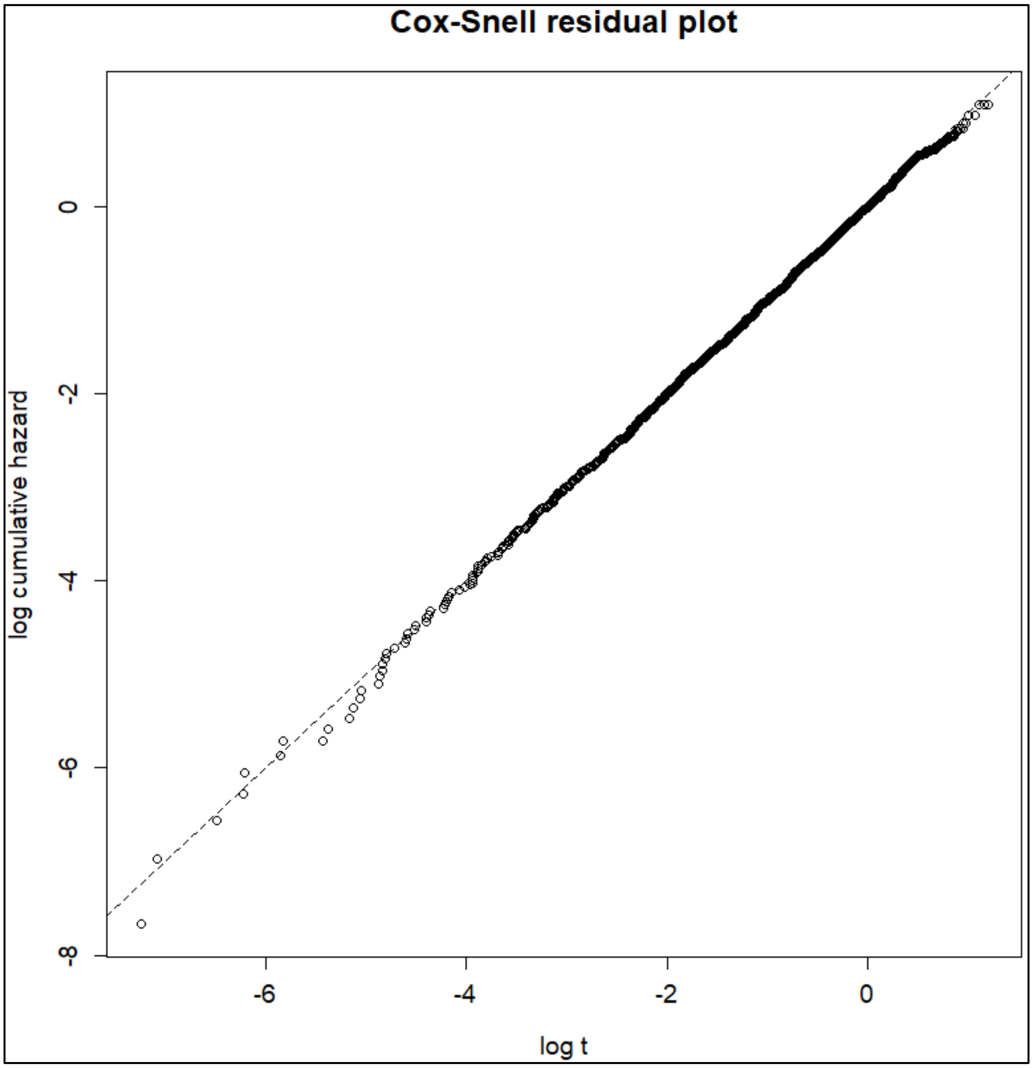

Figure 4: Cox-Snell Residuals for the hospitalization free survival

	chisq	df	p
factor(trt_ab)	1.243	1	0.265
factor(sex)	0.111	1	0.740
age	5.366	1	0.021
factor(etiology)	0.439	1	0.508
factor(afibflut)	2.513	1	0.113
factor(nationality)	5.364	2	0.068
bmi	3.739	1	0.053
cpxdur	0.896	1	0.344
GLOBAL	19.353	9	0.022

Figure 5: Schoenfeld Residuals Chi-Square test for proportionality assumption

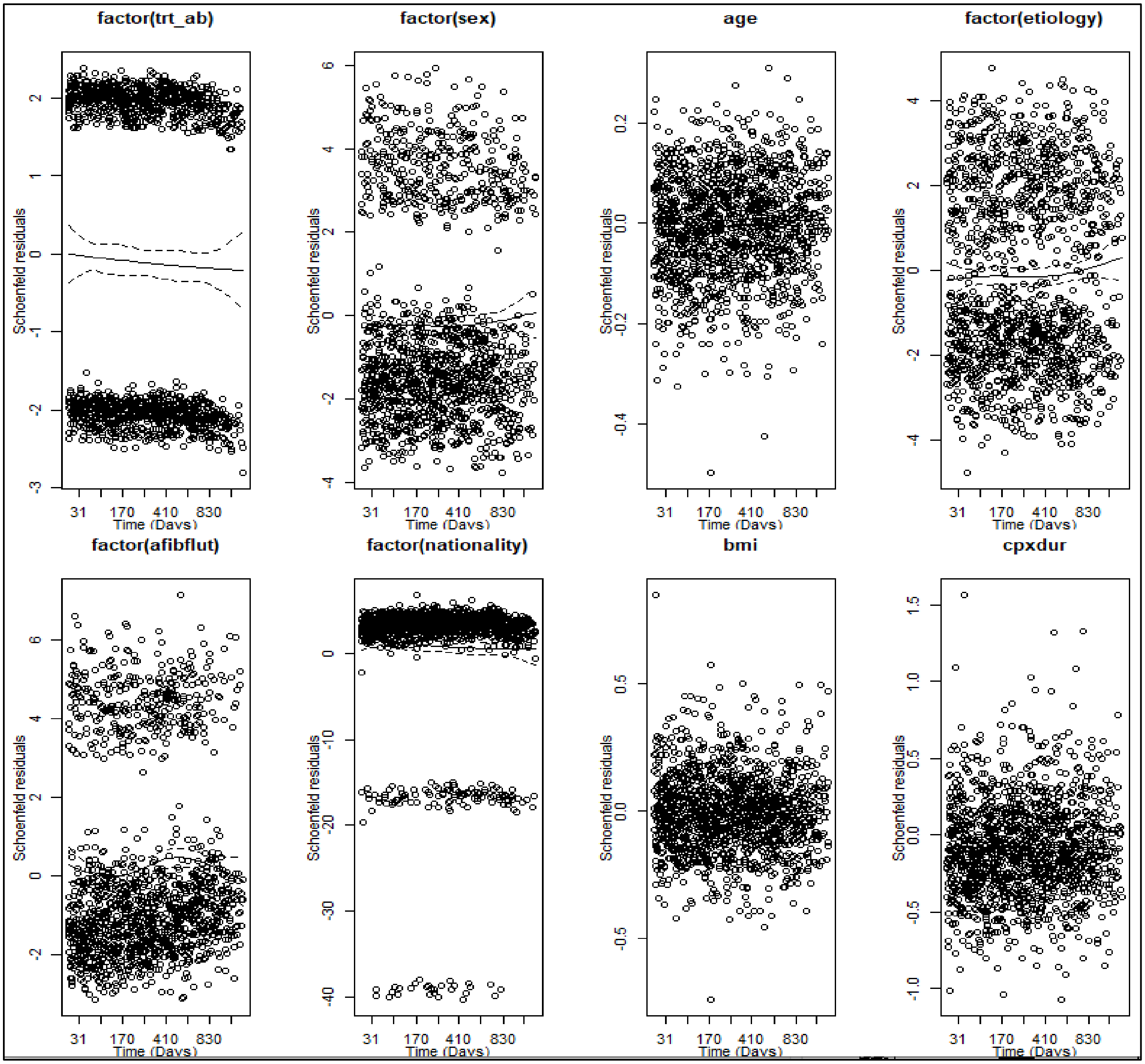

Figure 6: Schoenfeld Residuals plots for all the covariate

To further investigate the functional forms of the given covariate, we also evaluate the Martingale residuals for the continuous variables (Figure 7). The plots show that there is significant trend in the residuals for Age covariate suggesting that the current functional form of the Age covariate is not justified. A similar conclusion can be drawn for BMI and for the duration of cardiopulmonary exercise. We thus convert the 'Age' covariate and the 'BMI' covariate into categorical format and re do the whole analysis. The 'Age; covariate is converted into the following categories: category 1- for 'Age' less than 40 years, category 2- for 'Age' between 40-60 years and category 3- for 'Age' greater than 60 years. The BMI was categorized as less than 30 for fit people, 30-50 for moderately unfit people and greater than 50 for highly unhealthy people. The Cox-PH model was then fit again, and the results were produced (as shown in Figure 8). The significance of the 'Age' covariate is improved after categorizing it, however the significance of the BMI variable is reduced significantly and thus it becomes important to perform the Wald test (Table 2) for these covariates to identify their significance.

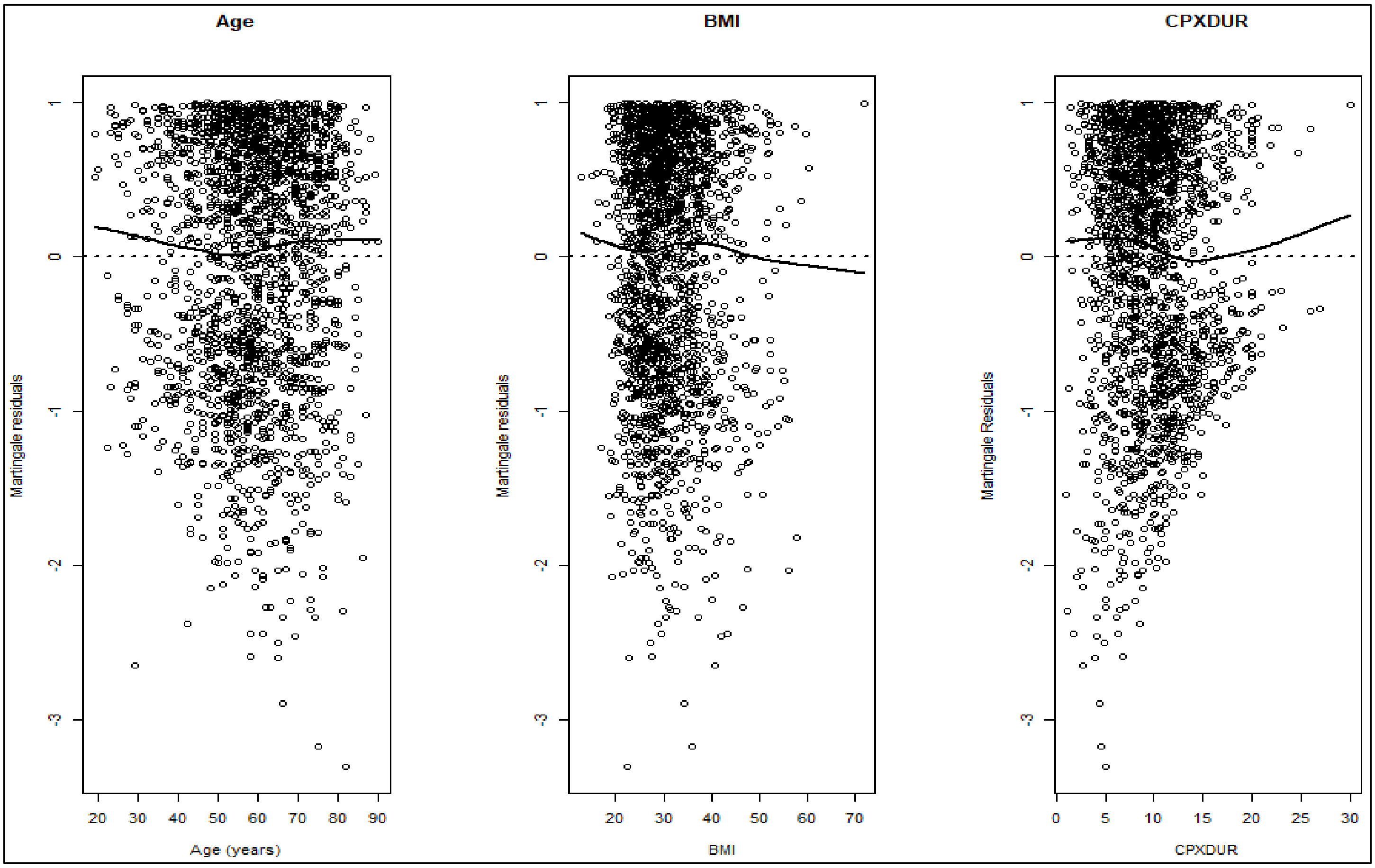

Figure 7: Martingale Residual for continuous variables

```
call:
coxph(formula = Surv(time, ttfe_status) ~ factor(trt_ab) + factor(sex) +
    factor(age_categorized) + factor(etiology) + factor(afibflut) +
    factor(nationality) + factor(bmi_categorized) + cpxdur, data = df_unique_ss1
_interpolated)

  n= 2130, number of events= 1419

                              coef exp(coef)  se(coef)       z Pr(>|z|)
factor(trt_ab)1          -0.106070  0.899362  0.053391  -1.987 0.046962 *
factor(sex)2             -0.218367  0.803831  0.064276  -3.397 0.000680 ***
factor(age_categorized)1 -0.250438  0.778460  0.111481  -2.246 0.024674 *
factor(age_categorized)2 -0.206513  0.813416  0.119152  -1.733 0.083062 .
factor(etiology)2        -0.109078  0.896661  0.058181  -1.875 0.060820 .
factor(afibflut)1         0.189872  1.209095  0.065328   2.906 0.003656 **
factor(nationality)2     -0.330973  0.718225  0.101739  -3.253 0.001141 **
factor(nationality)3     -0.651680  0.521169  0.192961  -3.377 0.000732 ***
factor(bmi_categorized)1 -0.006277  0.993743  0.056154  -0.112 0.911000
factor(bmi_categorized)2 -0.301424  0.739764  0.204245  -1.476 0.139999
cpxdur                   -0.092380  0.911759  0.007886 -11.715  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                         exp(coef) exp(-coef) lower .95 upper .95
factor(trt_ab)1             0.8994     1.1119    0.8100    0.9986
factor(sex)2                0.8038     1.2440    0.7087    0.9118
factor(age_categorized)1    0.7785     1.2846    0.6257    0.9686
factor(age_categorized)2    0.8134     1.2294    0.6440    1.0274
factor(etiology)2           0.8967     1.1152    0.8000    1.0050
factor(afibflut)1           1.2091     0.8271    1.0638    1.3743
factor(nationality)2        0.7182     1.3923    0.5884    0.8767
factor(nationality)3        0.5212     1.9188    0.3571    0.7607
factor(bmi_categorized)1    0.9937     1.0063    0.8902    1.1094
factor(bmi_categorized)2    0.7398     1.3518    0.4957    1.1039
cpxdur                      0.9118     1.0968    0.8978    0.9260

Concordance= 0.621  (se = 0.008 )
Likelihood ratio test= 228  on 11 df,   p=<2e-16
Wald test            = 219.7  on 11 df,   p=<2e-16
Score (logrank) test = 221.3  on 11 df,   p=<2e-16
```

Figure 8: Cox-PH model after transforming the “Age” and “BMI” covariate

For the Wald test (Table 2) it turns out that for 90% significance value the ‘Age’ and ‘Nationality’ covariate are significant but the categorical ‘BMI’ is not and thus it would be better to remove the BMI covariate for further analysis. So, the Cox-PH model was again fit and reported in Figure 9. However, as the effect of Nationality on the study is not too detrimental and as the nationality covariate is a bit un-balanced it is considered useful to stratify the model on Nationality while refitting. The Cox-Snell residuals produced after the analysis (Figure 10) shows that there is still a significant fit of the given model for the data. The Schoenfeld residuals produced in Figure 11 and 12 also shows that *the proportionality assumption is not violated now as the global proportionality statistics is now insignificant (as it is larger than 0.1) after the transformations. The p-value of all other covariates are also larger than 0.1 as shown in Figure 11 which represents improved significance for proportionality assumption*. The hazards for hospitalizations due to unit increment for various covariates are also summarized in Table 3 and are interpreted below

Table 2: Wald test for Categorical variable

Covariate	Chi-Sqare Statistics	P-Value
Age	5.23	0.07
BMI	2.21	0.331
Nationality	21.2	2.47E-05

```
Call:
coxph(formula = Surv(time, ttfe_status) ~ strata(factor(nationality)) +
    factor(trt_ab) + factor(sex_reorder) + factor(age_categorized) +
    factor(etiology_reorder) + factor(afibflut) + cpxdur, data = df_unique_ss1_i
nterpolated)

  n= 2130, number of events= 1419

                              coef exp(coef)  se(coef)       z Pr(>|z|)    
factor(trt_ab)1          -0.104804  0.900501  0.053338  -1.965 0.049424 *  
factor(sex_reorder)1      0.217884  1.243443  0.064308   3.388 0.000704 ***
factor(age_categorized)1 -0.225223  0.798338  0.110262  -2.043 0.041090 *  
factor(age_categorized)2 -0.169590  0.844011  0.115354  -1.470 0.141515    
factor(etiology_reorder)1  0.110382  1.116705  0.058189   1.897 0.057833 .  
factor(afibflut)1         0.191469  1.211028  0.065280   2.933 0.003357 ** 
cpxdur                   -0.090125  0.913817  0.007611 -11.842  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                         exp(coef) exp(-coef) lower .95 upper .95
factor(trt_ab)1             0.9005     1.1105    0.8111    0.9997
factor(sex_reorder)1        1.2434     0.8042    1.0962    1.4105
factor(age_categorized)1    0.7983     1.2526    0.6432    0.9909
factor(age_categorized)2    0.8440     1.1848    0.6732    1.0581
factor(etiology_reorder)1   1.1167     0.8955    0.9963    1.2516
factor(afibflut)1           1.2110     0.8257    1.0656    1.3763
cpxdur                      0.9138     1.0943    0.9003    0.9276

Concordance= 0.614  (se = 0.008 )
Likelihood ratio test= 204.2  on 7 df,   p=<2e-16
Wald test            = 197.2  on 7 df,   p=<2e-16
Score (logrank) test = 198.1  on 7 df,   p=<2e-16
```

Figure 9: Cox-PH model results after data transformation and BMI covariate removal

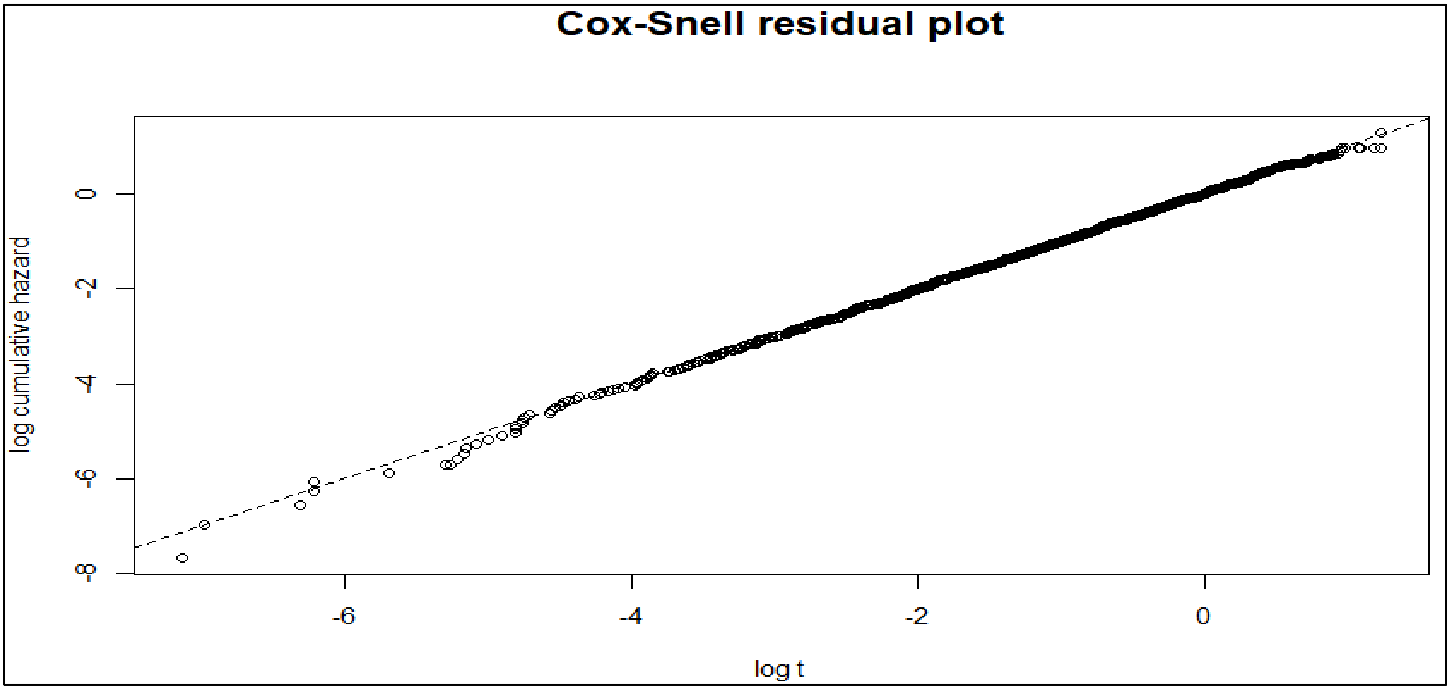

Figure 10: Cox-Snell residuals after transforming

```
> sch = cox.zph(obj.trans.drop)
> print(sch)
                              chisq df    p
factor(trt_ab)                1.358  1 0.24
factor(sex_reorder)           0.240  1 0.62
factor(age_categorized)       2.636  2 0.27
factor(etiology_reorder)      0.683  1 0.41
factor(afibflut)              2.180  1 0.14
cpxdur                        0.884  1 0.35
GLOBAL                       10.445  7 0.16
```

Figure 11: Schoenfeld residuals Chi-Square test for proportionality assumption

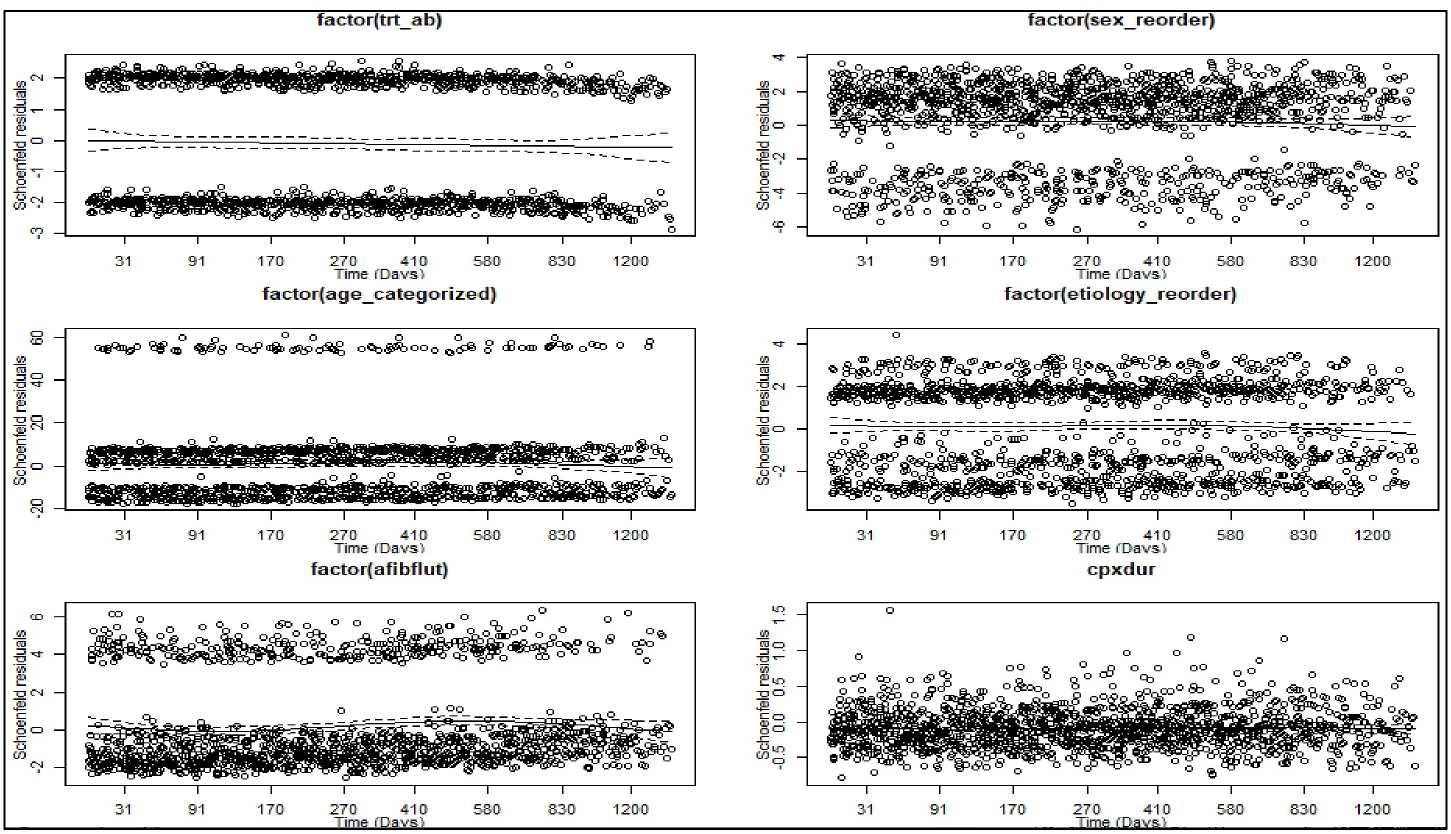

Figure 12: Schoenfeld Residuals plots for all the covariate after covariate processing

Table 3: Final Hazard Ratios of covariates for time to hospitalization

	Coefficient	Hazard Ratio	Standard Error	P-Value
Treatment: Exercise (1) vs Basic Care (0)	-0.105	0.901	0.053	0.049
Sex: Male(1) vs Female (0)	0.218	1.243	0.064	0.001
Age category: 40-60 yrs (1) vs <40 yrs (0)	-0.225	0.798	0.110	0.041
Age category: >60 yrs (2) vs <40 yrs (0)	-0.170	0.844	0.115	0.142
Etiology: Ischemic (1) vs Non-Ischemic (0)	0.110	1.117	0.058	0.058
History of atrial fibrillation: Yes(1) vs No(0)	0.191	1.211	0.065	0.003
Duration of Cardipulmonary Exercise test (mins)	-0.090	0.914	0.008	< 2e-16

From Table 3 it can be identified that aerobic exercise helps to reduce the hazard for hospitalization by around 10%. It can also be seen that the males are at a higher risk of getting hospitalized as compared to the females. From the analysis it can also be shown that the effect of exercise on elder people is much more advantageous than for the younger people as the hazard for the elderly people is reduced as compared to the younger people. Subjects having a history of ischemic etiology and atrial fibrillation are having a higher hazard for being hospitalized as compared to the subjects which are free from any of these histories.

Pocock's Proportional Win Fraction Model

We analyze the data using the Pocock's proportional win-fractions (PW) model by prioritizing death over hospitalization. The results obtained are shown in Figure 13 and the output consists of three parts. The first part presents some descriptive statistics on the proportions of win-loss status among all (n2) = 2,267,385 pairs. According to the output, 9.6% of them are determined by death; 73.4% by hospitalization, and the remaining 16.9% are indeterminate. It also reports an overall (Wald) test with p-value ~ 0, suggesting that, at the conventional 0.05 level, the 7 covariates are significantly associated with the composite outcome. The second part presents a table for the estimates and standard errors of the regression coefficient, along with their corresponding p -value for testing the coefficient being zero. The third and the most informative part, tabulates the estimated win ratios (exponential of the regression coefficients) and their associated 95% confidence intervals. We can see that using the current model a patient in provided with the aerobic exercise is 7% more likely to have a better priority-adjusted composite outcome for survival time than the one in usual care. This p-value indicates that this difference is presently statistically in-significant. We also generate the plots for the scores of the covariates as obtained by the present model in Figure 14. By observing the scores, it can be said that there is a high amount of trend present for the 'Age' covariate thus violating the proportionality assumption as already considered in the previous analysis. To improve the same and to improve the interpretability the 'Nationality' and the 'Age' covariate were transformed. The 'Nationality' covariate was converted into two other indicator covariates which are used to indicates 1) if a subject is from USA or not and 2) if a subject is from Canada or not. The Age covariate was also replaced by two indicator columns which indicate 1) if the subject's age is less than 40 years or not and 2) if the subject is elder than 60 years or not. By generating the appropriate transformations, the Pocock's Win Ratio test was again conducted, the results are reported in Figure 15 and are tabulated in Table 4.

```
Call:
pwreg(time = time_, status = status_, Z = Z_, ID = ID_)

Proportional win-fractions regression models for priority-adjusted composite endpoint

     (Mao and Wang, 2020+, under review):

Total number of pairs: 2267385
Wins-losses on death:  218687 (9.6%)
Wins-losses on non-fatal event:  1664992 (73.4%)
Indeterminate pairs 383706 (16.9%)

Newton-Raphson algorithm converged in 5 iterations.

Overall test: chisq test with 7 degrees of freedom;
 Wald statistic 211.5 with p-value 0

Estimates for Regression parameters:

                     Estimate          se z.value   p.value
age                 0.0018983  0.0025550  0.7430  0.457496
trt_ab              0.0647851  0.0574599  1.1275  0.259538
nationality         0.3371271  0.0749317  4.4991 6.823e-06 ***
cpxdur              0.1083622  0.0091845 11.7984 < 2.2e-16 ***
afibflut           -0.2216538  0.0699945 -3.1667  0.001542 **
etiology_converted -0.1423215  0.0633694 -2.2459  0.024710 *
sex_converted      -0.3310406  0.0684329 -4.8374 1.315e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Point and interval estimates for for win ratios:

                   Win Ratio 95% lower CL 95% higher CL
age                1.0019001    0.9968955     1.0069298
trt_ab             1.0669298    0.9532919     1.1941139
nationality        1.4009172    1.2095685     1.6225363
cpxdur             1.1144514    1.0945693     1.1346946
afibflut           0.8011927    0.6984858     0.9190020
etiology_converted 0.8673424    0.7660382     0.9820434
sex_converted      0.7181760    0.6280305     0.8212608
```

Figure 13: Pocock's Proportional Win Ratio test

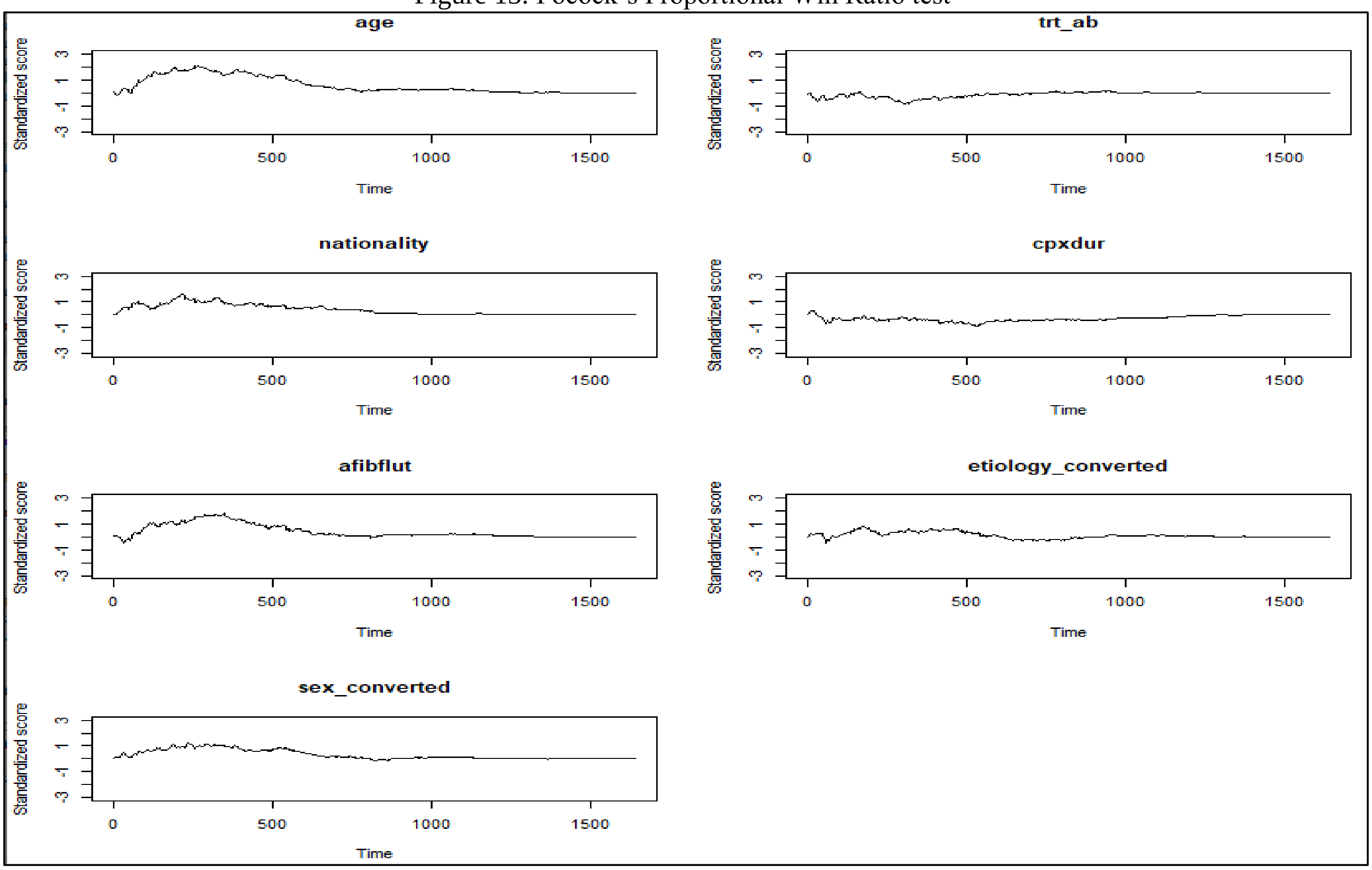

Figure 14: Scores generated after fitting the Pocock's Proportional Win Ratio test

```
Call:
pwreg(time = time__, status = status__, Z = Z__, ID = ID__)

Proportional win-fractions regression models for priority-adjusted composite endpoin
t

     (Mao and Wang, 2020+, under review):

Total number of pairs: 2267385
Wins-losses on death:  218687 (9.6%)
Wins-losses on non-fatal event:  1664992 (73.4%)
Indeterminate pairs 383706 (16.9%)

Newton-Raphson algorithm converged in 5 iterations.

Overall test: chisq test with 9 degrees of freedom;
 Wald statistic 212.2 with p-value 0

Estimates for Regression parameters:

                                    Estimate         se z.value   p.value
trt_ab                             0.0597795  0.0575392  1.0389 0.2988344
cpxdur                             0.1074687  0.0091411 11.7567 < 2.2e-16 ***
afibflut                          -0.2198948  0.0697189 -3.1540 0.0016104 **
etiology_converted                -0.1443846  0.0625375 -2.3088 0.0209565 *
sex_converted                     -0.3289699  0.0686779 -4.7900 1.668e-06 ***
nationality_USA_versus_others     -0.6760920  0.1993712 -3.3911 0.0006961 ***
nationality_Canada_versus_others  -0.3591487  0.2237188 -1.6054 0.1084151
Age_less_thn_40_vs_others         -0.2773005  0.1171898 -2.3663 0.0179693 *
Age_more_thn_60_vs_others         -0.0625201  0.0622913 -1.0037 0.3155366
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Point and interval estimates for for win ratios:

                                  Win Ratio 95% lower CL 95% higher CL
trt_ab                            1.0616025    0.9483846     1.1883363
cpxdur                            1.1134560    1.0936848     1.1335846
afibflut                          0.8026032    0.7000935     0.9201227
etiology_converted                0.8655548    0.7657068     0.9784229
sex_converted                     0.7196647    0.6290301     0.8233585
nationality_USA_versus_others     0.5086007    0.3440898     0.7517649
nationality_Canada_versus_others  0.6982705    0.4503953     1.0825637
Age_less_thn_40_vs_others         0.7578267    0.6023068     0.9535031
Age_more_thn_60_vs_others         0.9393942    0.8314295     1.0613786
```

Figure 15: Pocock’s Proportional Win Ratio test after transformations

Table 4: Pocock’s Proportional Win Ratio multiple regression analysis for HF-Action data after covariate transformation

	Win Ratio	**95% lower CL**	**95% higher CL**	**P-Values**
Sex: Male(1) vs Female (0)	0.720	0.629	0.823	1.67E-06
Treatment: Exercise (1) vs Basic Care (0)	1.062	0.948	1.188	0.2988
Duration of Cardipulmonary Exercise test (mins)	1.113	1.094	1.134	< 2.2E-16
History of atrial fibrillation: Yes(1) vs No(0)	0.803	0.700	0.920	0.0016
Etiology: Ischemic (1) vs Non-Ischemic (0)	0.866	0.766	0.978	0.0210
Nationality: USA (1) vs others (0)	0.509	0.344	0.752	0.0007
Nationality: Canada (1) vs others (0)	0.698	0.450	1.083	0.1084
Age: less than 40 yrs(1) vs greater than 40 yrs (0)	0.758	0.602	0.954	0.0180
Age: more than 60 yrs(1) vs less than 60 yrs (0)	0.939	0.831	1.061	0.3155

From table 4 it can be seen that, the inferences drawn from the Pocock’s Proportional Wins (PW) model are similar to those drawn from the Cox-PH model when it comes to analyze the effect of aerobic exercise versus basic care. However, for the PW model the difference is in-significant. The inferences for History of atrial fibrillation and presence of ischemic etiology are also consistent between the PW and Cox-PH model, where the subjects with the history are more likely to have the priority adjusted outcome (death) as compared to the ones with any history. It can also be noted that the female subjects are more likely to have a better (priority-adjusted) outcome than male subjects (they are 28% less likely to die as compared to male subjects), which is still consistent with the earlier findings where the hazard for male subjects was higher than the female ones for time to first event. For Age covariate as well, we find a similar inference as before where the younger (<40 years) subjects have a 25% higher risk as compared to others. As the Cox-PH model was stratified on Nationality there was no inference available for the hazard contributed by it. However, for the PW model we find that the subjects belonging to USA are 24.2% time more likely to the priority adjusted outcome as compared to others, while those subjects in Canada are just 6% more likely than others. In order to test the significance of Nationality and Age covariate, Wald statistics were generated (Table 5) as done in (Table 2). Comparing both the tables the same conclusion is drawn that on a significance level of 90% both covariates demonstrate a significant level of difference in survival probabilities of subjects. The score plots (Figure 16) are generated for each covariate which now demonstrate that there are still some trends present for the Age indicator of more than 60 years and also for the status indicator for history of atrial fibrillation. However, barring them all other covariates are nor well behaved.

Table 5: Wald test for

Covariate	Chi-Sqare Statistics	P-Value
Age	5.95	0.051
Nationality	19.3	6.57E-05

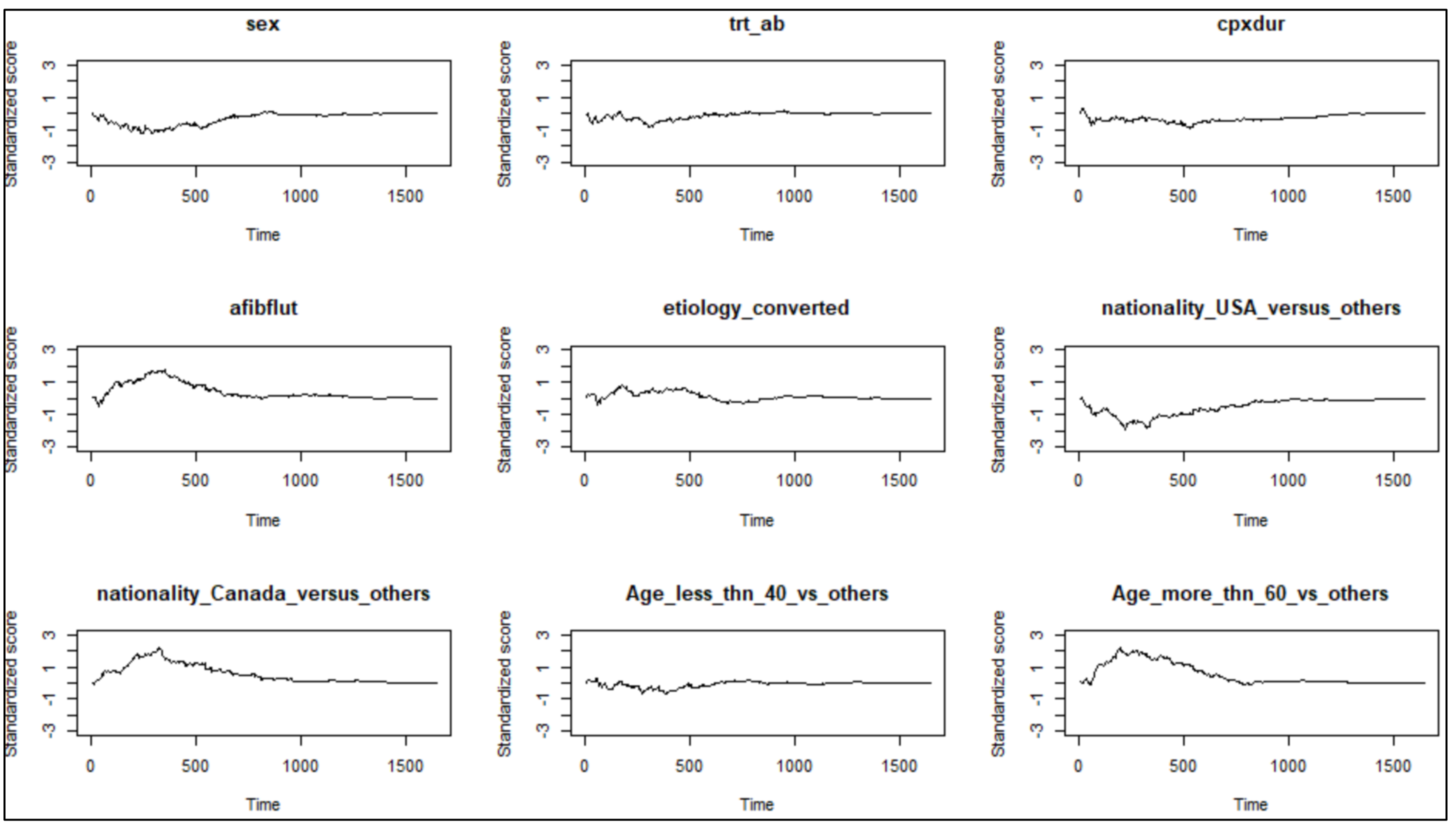

Figure 16: Scores generated for Pocock's Proportional Win Ratio test with transformed Age and Nationality covariates

Supplementary Studies (Cox-PH model for time to death)

To further analyze the given data a Cox-PH model was fit for the events corresponding to terminal event as death or right censoring by performing appropriate interpolation for the missing values as discussed in earlier. The results obtained by fitting the Cox-PH model for death as terminal events are demonstrated in Figure 17 and Table 6. It can be seen that the group with treatment have a reduction in hazard risk for death due to Aerobic Exercise however the effect of the Exercise is not significant. Apart from that the Male members having an history of ischemic etiology and atrial fibrillation have a very high hazard as opposed to others. For comparison in Nationality people belonging to Canada have a higher risk than to those belonging to USA, however people of France have a significantly lower risk than people of USA. The baseline hazard for the model is shown in Figure 18. Figure 19 shows the Cox-Snell residuals which demonstrate a good level of fit for the model. The proportionality assumption is also kept intact for the model as the global chi-square statistics generated by Schoenfeld residuals is insignificant (p-value 0.43) as shown in Figure 20. Figure 21 and 22 shows the plot for the Schoenfeld and Martingale residuals.

```
Call:
coxph(formula = Surv(time, status) ~ factor(trt_ab) + factor(sex_rearranged) +
    age + factor(etiology_converted) + factor(afibflut) + factor(nationality) +
    cpxdur, data = data_1.end.event.interpolated)

  n= 2130, number of events= 351

                                coef exp(coef)  se(coef)       z Pr(>|z|)
factor(trt_ab)1            -0.065367  0.936723  0.107026  -0.611   0.5414
factor(sex_rearranged)1     0.618139  1.855472  0.140210   4.409 1.04e-05 ***
age                         0.009431  1.009476  0.004883   1.932   0.0534 .
factor(etiology_converted)1 0.063941  1.066029  0.116500   0.549   0.5831
factor(afibflut)1           0.311478  1.365442  0.119758   2.601   0.0093 **
factor(nationality)1        0.108755  1.114889  0.189597   0.574   0.5662
factor(nationality)2       -0.962834  0.381809  0.581262  -1.656   0.0976 .
cpxdur                     -0.172621  0.841456  0.016998 -10.156  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                            exp(coef) exp(-coef) lower .95 upper .95
factor(trt_ab)1                0.9367     1.0676    0.7595     1.155
factor(sex_rearranged)1        1.8555     0.5389    1.4096     2.442
age                            1.0095     0.9906    0.9999     1.019
factor(etiology_converted)1    1.0660     0.9381    0.8484     1.339
factor(afibflut)1              1.3654     0.7324    1.0798     1.727
factor(nationality)1           1.1149     0.8970    0.7689     1.617
factor(nationality)2           0.3818     2.6191    0.1222     1.193
cpxdur                         0.8415     1.1884    0.8139     0.870

Concordance= 0.707  (se = 0.014 )
Likelihood ratio test= 186.8  on 8 df,   p=<2e-16
Wald test            = 167.9  on 8 df,   p=<2e-16
Score (logrank) test = 173.2  on 8 df,   p=<2e-16
```

Figure 17: Proportional Hazard model for time to death

Table 6: Summarizing the hazard risk for various covariates for time to death

	Coefficient	Hazard Ratio	Standard Error	P-Value
Treatment: Exercise (1) vs Basic Care (0)	-0.0654	0.9367	0.1070	0.5414
Sex: Male(1) vs Female (0)	0.6181	1.8555	0.1402	1.04E-05
Age (Years)	0.0094	1.0095	0.0049	0.0534
Etiology: Ischemic (1) vs Non-Ischemic (0)	0.0639	1.0660	0.1165	0.5831
History of atrial fibrillation: Yes(1) vs No(0)	0.3115	1.3654	0.1198	0.0093
Nationality Canada (1) vs USA (0)	0.1088	1.1149	0.1896	0.5662
Nationality France (2) vs USA (0)	-0.9628	0.3818	0.5813	0.0976
Duration of Cardipulmonary Exercise test (mins)	-0.1726	0.8415	0.0170	2.00E-16

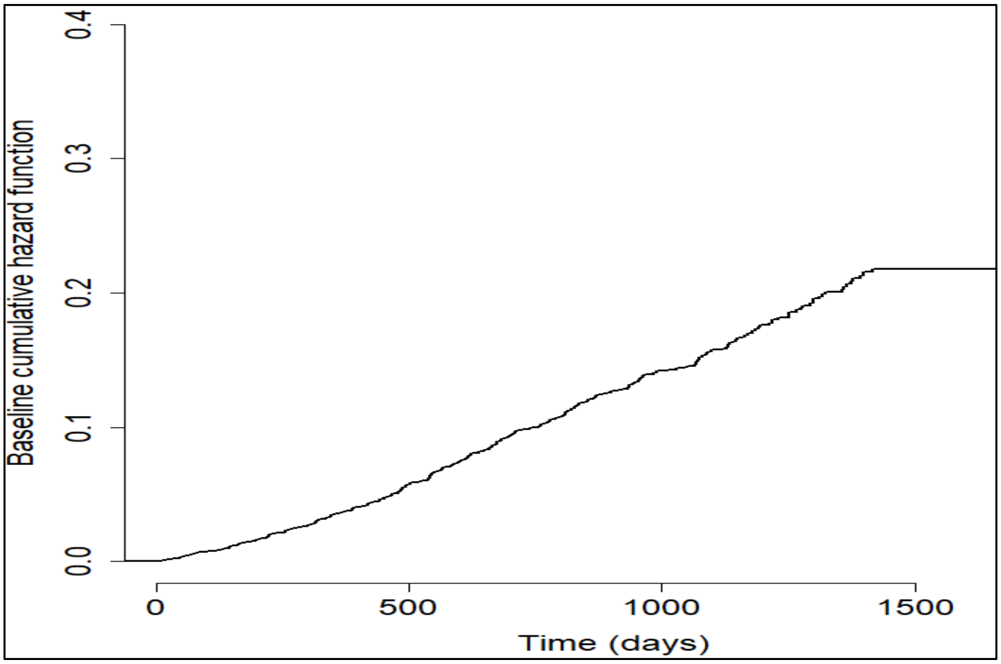

Figure 18: Baseline Hazard for the Cox-PH model for death

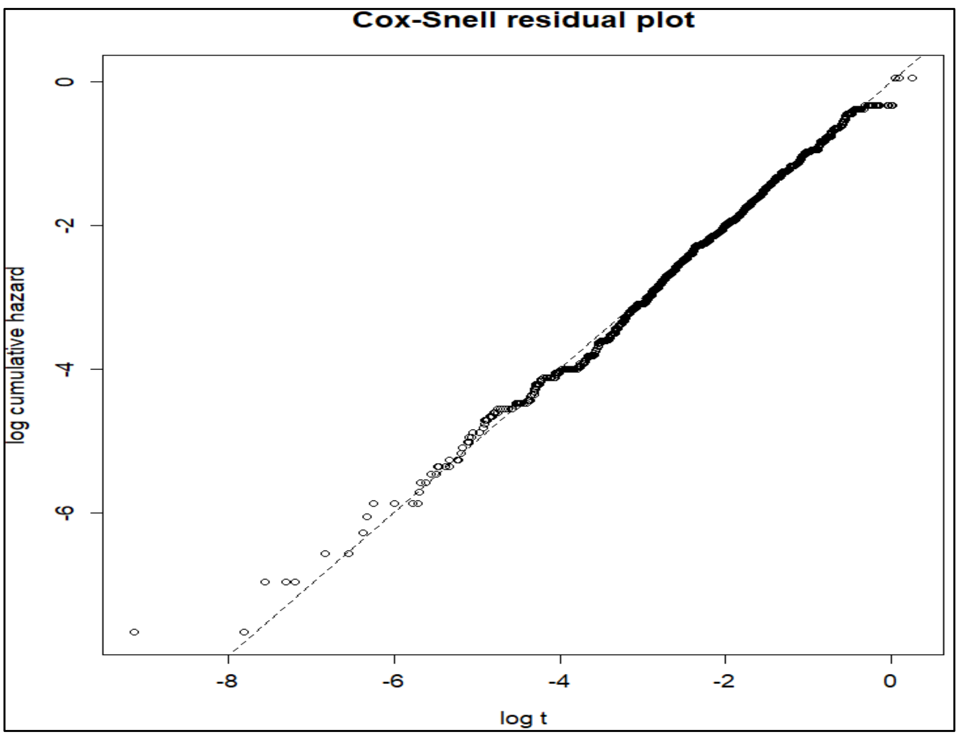

Figure 19: Cox-Snell Residual for the Cox-PH model

	chisq	df	p
factor(trt_ab)	1.8025	1	0.18
factor(sex_rearranged)	0.0695	1	0.79
age	1.1734	1	0.28
factor(etiology_converted)	1.8670	1	0.17
factor(afibflut)	0.8634	1	0.35
factor(nationality)	3.0077	2	0.22
cpxdur	0.0902	1	0.76
GLOBAL	8.0426	8	0.43

Figure 20: Schoenfeld residuals Chi-Square test for proportionality assumption

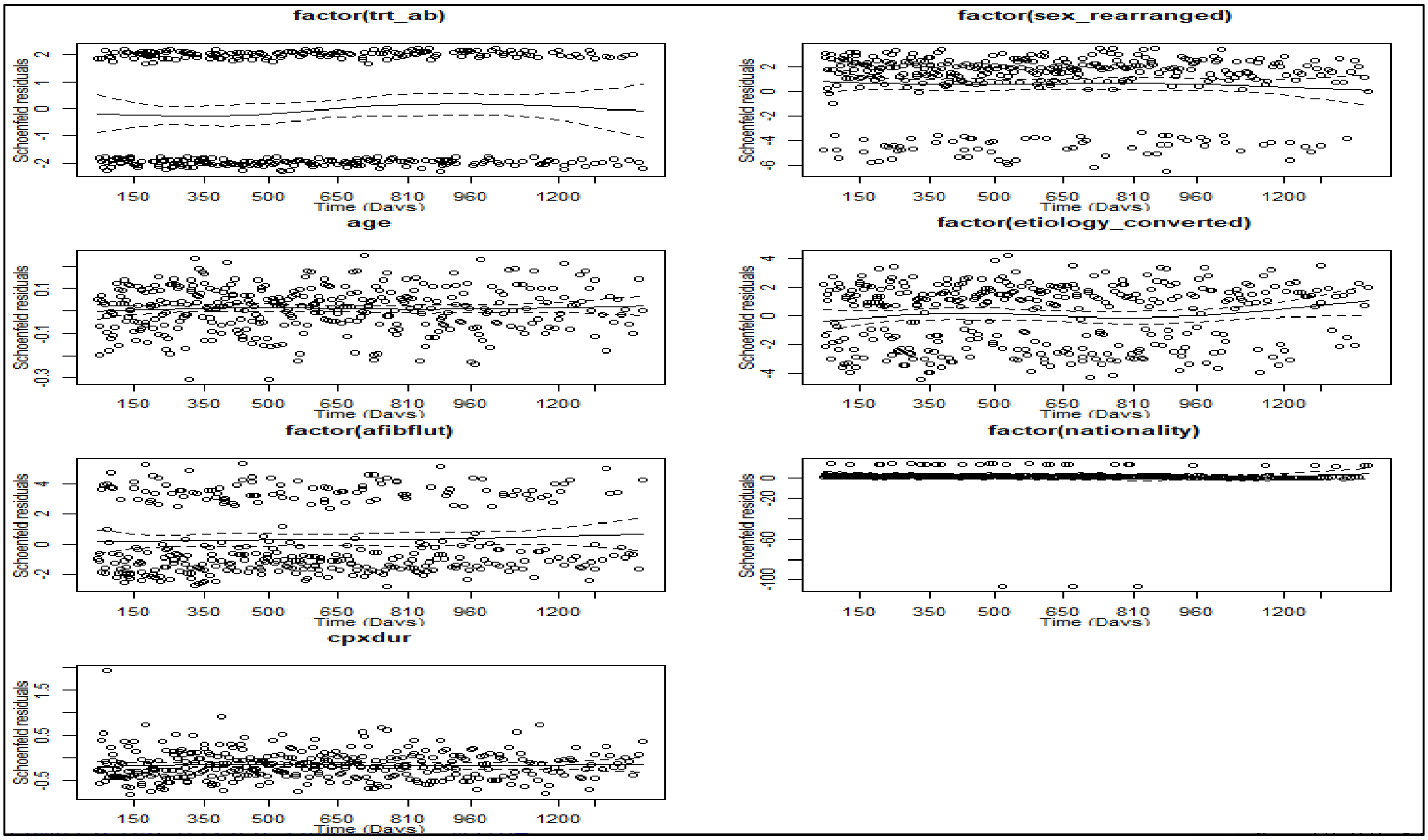

Figure 21: Schoenfeld residuals plots for the proportionality test

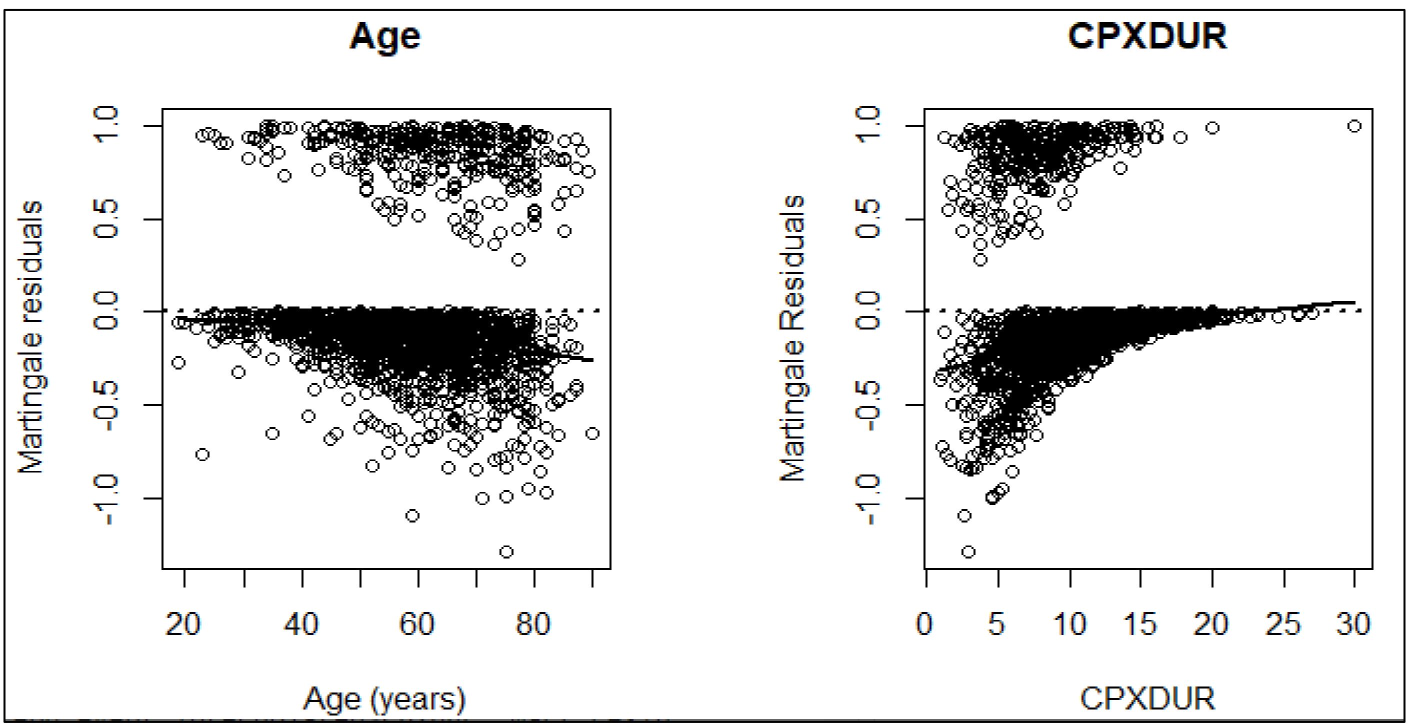

Figure 22: Martingale residual verifying the functional form for Age and duration of cardiopulmonary exercise test

Supplementary Studies (Proportional Means Model for Recurrent Hospitalization)

We also study the recurrent hospitalization events by fitting the proportional means models. The output of the same is demonstrated in Figure 23 which shows that mean rate of hospitalization is approximately the same for the subjects with aerobic exercise or usual care. The hazard rate for hospitalization increases for a USA male individual with a history of ischemic etiology and atrial fibrillation when compared to a female subject not residing in USA and not having a history of ischemic etiology or atrial fibrillation. Figure 24 demonstrate the base hazard for French females of 50 years of in Age having 13 minutes of duration for cardiopulmonary exercise test and having no history for ischemic etiology or atrial fibrillation have differences in their hazard rate for control and treatment group. Also, their hazard rate is lower when compared to the USA males of 50 years of in Age having 13 minutes of duration for cardiopulmonary exercise test having a history of ischemic etiology or atrial fibrillation.

```
Call:
coxph(formula = Surv(start, stop, status_recc) ~ factor(trt_ab) +
    factor(sex_rearranged) + age + cpxdur + factor(nationality) +
    factor(etiology_converted) + factor(afibflut), data = data_1.recc.hosp,
    cluster = patid)

  n= 6006, number of events= 4246
   (40 observations deleted due to missingness)

                                coef exp(coef)  se(coef) robust se       z Pr(>|z|)
factor(trt_ab)1            -0.002832  0.997172  0.030748  0.060417  -0.047 0.962611
factor(sex_rearranged)1     0.210620  1.234444  0.036563  0.073816   2.853 0.004326 **
age                        -0.008580  0.991457  0.001381  0.002924  -2.934 0.003341 **
cpxdur                     -0.104351  0.900909  0.004529  0.008928 -11.687  < 2e-16 ***
factor(nationality)2       -0.522005  0.593330  0.073430  0.109975  -4.747 2.07e-06 ***
factor(nationality)3       -0.460124  0.631206  0.119320  0.247152  -1.862 0.062645 .
factor(etiology_converted)1 0.195997  1.216523  0.033884  0.066287   2.957 0.003108 **
factor(afibflut)1           0.283563  1.327852  0.036228  0.074986   3.782 0.000156 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                           exp(coef) exp(-coef) lower .95 upper .95
factor(trt_ab)1               0.9972     1.0028    0.8858    1.1225
factor(sex_rearranged)1       1.2344     0.8101    1.0682    1.4266
age                           0.9915     1.0086    0.9858    0.9972
cpxdur                        0.9009     1.1100    0.8853    0.9168
factor(nationality)2          0.5933     1.6854    0.4783    0.7360
factor(nationality)3          0.6312     1.5843    0.3889    1.0246
factor(etiology_converted)1   1.2165     0.8220    1.0683    1.3853
factor(afibflut)1             1.3279     0.7531    1.1464    1.5381

Concordance= 0.63  (se = 0.008 )
Likelihood ratio test= 817.1  on 8 df,   p=<2e-16
Wald test            = 239.8  on 8 df,   p=<2e-16
Score (logrank) test = 781.3  on 8 df,   p=<2e-16,   Robust = 184.7  p=<2e-16
```

Figure 23: Proportional Means model for recurrent hospitalization

Table 7: Summarizing the hazard risk for Proportional Means models

	coefficient	hazard ratio	standard error	p-value
Treatment: Exercise (1) vs Basic Care (0)	-0.00300	0.99700	0.06000	0.96300
Sex: Male(1) vs Female (0)	0.21100	1.23491	0.07400	0.00400
Age (Years)	-0.00900	0.99104	0.00300	0.00300
Duration of Cardipulmonary Exercise test (mins)	-0.10400	0.90123	0.00900	0.00000
Nationality Canada (2) vs USA (1)	-0.52200	0.59333	0.11000	0.00000
Nationality France (3) vs USA (1)	-0.46000	0.63128	0.24700	0.06300
Etiology: Ischemic (1) vs Non-Ischemic (0)	0.19600	1.21652	0.06629	0.00311
History of atrial fibrillation: Yes(1) vs No(0)	0.28356	1.32785	0.07499	0.00016

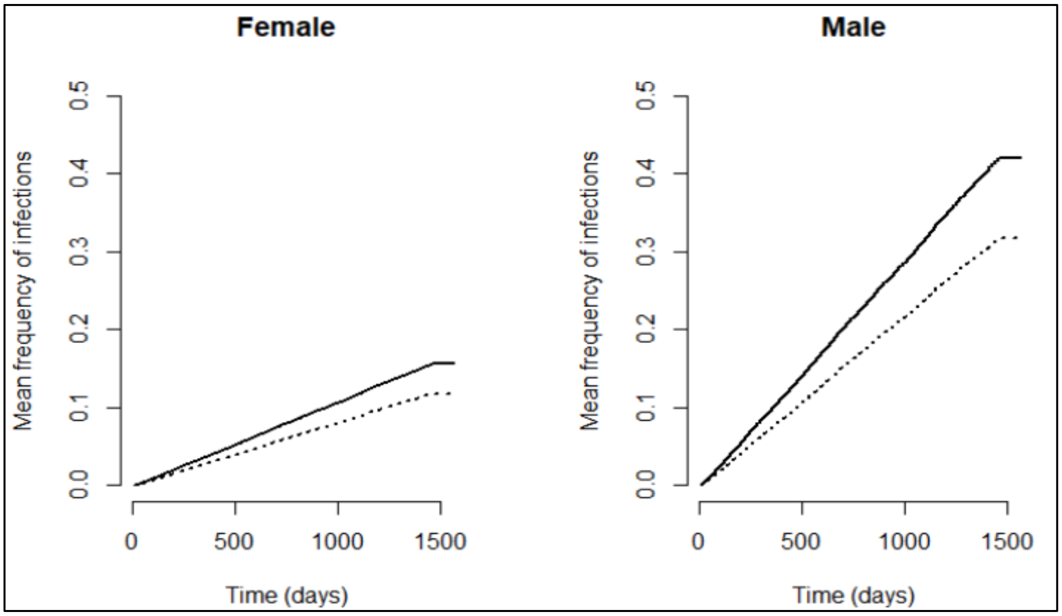

Figure 24: Proportional Means model for recurrent hospitalization