Name: Bunyod Tusmatov

Date: 12/02/2018

Biostatistics Project: *this project reviews and re-implements a published paper in biostatistics.*

Paper used: ***"****Do Nationally Representative Cutpoints for Clinical Muscle Weakness Predict Mortality? Results From 9 Years of Follow-up in the Health and Retirement Study" by Kate Duchowny*

Summary of the Study

The paper examines the association between muscle weakness, as measured by handgrip strength, and mortality of elderly people. The paper is based on 9-year observational study from the Health and Retirement Study (HRS) which is the longest running longitudinal study of older Americans in the United States. HRS started in 1992 and randomly selected participants and interviewed biannually since 1992. In 2006, half the sample of HRS participants was randomly selected for an enhanced face-to-face interview that included physical measurement. During the physical measurement hand grip strength and other body measurements were measured using a tool. Baseline collection of variables began in 2006 and was repeated every 4 years going forward. The study uses sample of 8,326 of older adults (≥ 65 yrs) from the Health and Retirement Study (HRS).

The study is 9-year observational study. In this study, time-varying clinical muscle weakness, as defined by handgrip strength cutpoints (see Appendix Table A2), was the primary exposure and time to death was the outcome of interest. The paper uses t-test and chi-square tests for analyzing bivariate differences between weak versus non-weak individuals at the baseline. The Kaplan-Meier method was used to generate time-to-event unadjusted baseline survival curves and log-rank test was used for testing the difference. Moreover, the Cox proportional hazards regression model used to estimate hazard ratios (HRs), while adjusting for covariates, for mortality among weak versus non-weak individuals. The study includes several covariates such as age, sex, race, education, physical activity, etc.

The paper finds that weak individuals had a steeper decline in their survival trajectory, compared to non-weak individuals. After adjusting for covariates, weak individuals were over 50% more likely to die earlier than non-work individuals. The paper used SAS for all the tests.

As mentioned above, the study used Kaplan-Meier survival curves to see difference in mortality between weak and non-weak people. Moreover, the study used log-rank tests to test weather survival curves different or not. The Kaplan-Meier survival curve is very good way of visualizing how subjects surviving for a certain period of time and how groups differ in terms of their survival. However, one should be aware of two most important assumptions that the Kaplan-Meier method makes. The first assumption it makes is that subjects who are censored at any time have the same survival prospects as those who continue to be followed. This assumption should be a concern where large number of people are censored during the study and we don't know what happened to them. However, according to the article only 23 people were censored because of follow up and remaining 6,570 (*79%*) were censored at the end of study. Therefore, the first assumption should not be a concern.

The second assumption it makes is the survival probabilities are the same for subjects recruited early and late in the study. In this study, a randomly chosen half of the participants participated in the interview in 2006 and another half were interviewed in 2008. After the first interview, they were interviewed in every four years. In order to represent American senior population new participants were added in 2010, 2012 and 2014. Although the number of new participants is relatively small, it is incorrect to assume that new participants have the same probability of survival. For instance, a new participant who joined at age 65 in

2014 were censored at the end of 2014(cutoff point), and he/she is very likely to survive by the end by the end of the year compare to a participant who joined at age 65 in 2006 who is already 73 years old. Moreover, new participants are more likely to be physically strong and they might influence the non-weak curve. Therefore, I think that the second assumption might have been violated in this study. One approach that could solve this issue is to focus only on 2006 participants since it is a large enough sample and view their survival curve.

Another method that is used in the study is adjusted Cox model for time-dependent variables were used to estimate hazard ratio (HRs) for the event of interest(mortality) among weak and non-weak participants. The Cox model is a great model to examine if we are interested in adjusted survival curves. Sometimes we are interested in multiple-covariate adjusted curves and Cox proportional model is a great semi-parametric method for estimating survival curves while adjusting for covariates.

Since Cox model is semiparametric (contains parametric and non-parametric components), it is robust when the baseline hazard is incorrect or misspecified because the baseline hazard function doesn't have to be estimated in order to estimate hazard ration. The Cox model assumes that there is a constant relationship between dependent and covariate variables. This implies that if person X had a risk of mortality 5 times higher than person Y, later time this difference will stay proportionally the same. In this study, this assumption is little concerning as older people's risk of mortality fluctuates quickly, and their proportional difference might change rapidly. In addition to this, the study converted muscle strength into binary variable. Although this binary variable might change in later years, this model doesn't encounter possibilities of smaller changes. For instance, if someone had grip strength 50 and later years in changed to 36. This person lost a lot of muscle strength but the model will notice this difference. Therefore, a new model should resolve this issue.

Table 1: T-test at the baseline with new data set

	Weak (*n=3290*) *Mean*	Non-Weak (*n=3514*) *Mean*	p-value
Age(y)	85.20	77.66	2.2e-16
BMI	29.15	30.18	2.274e-12

The paper provides a table that includes baseline demographic characteristics of the study. I have calculated similar demographic characteristics (**Table 1**) with my data set and compared to the table in the study paper. All the characteristics are very similar, which allows me to replicate unadjusted K-M and fully-adjusted Cox models (weighted).

In order to replicate the K-M method, I made a few assumptions related to organizing the data for the survival object in R. The study used two cohorts that were interviewed in 2006 and 2008 (no overlap of subjects). The second interview for the cohort 2006 was interviewed again in 2010 and 2014. The second interview for the cohort 2008 was in 2012. Moreover, new subjects were added in 2010, 2012, and 2014. In order to organize this data, I created time1 variable that indicated start data and time2 indicated end of interval or death year of a subject, whichever came first (see Appendix: Figure A1). After fitting the K-M survival curve, a similar survival curves for weak and non-weak groups has been obtained (see Figure 1). It is clear that two curves are clearly different from each other given we are working with a large sample size.

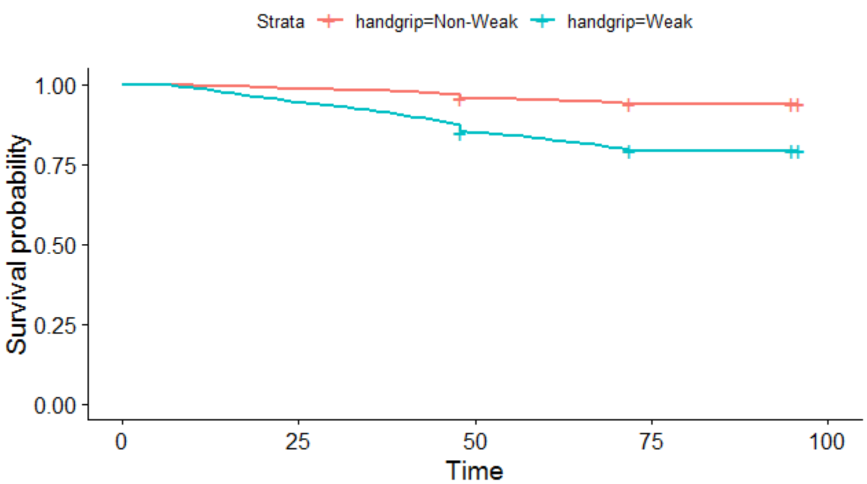

Table 2: Replicated Two Models

	HR	lower CI	upper CI
Model 2	HR	lower CI	upper CI
Weak[a]	2.0238***	1.7126	2.3914
AGE	1.0874***	1.0790	1.0960
GENDER[b]	0.5978***	0.5248	0.6809
Education[c]			
GED	1.8080***	1.3723	2.3820
High School	1.3778***	1.1711	1.6210
Some college	1.1748	0.8182	1.6867
RACE[d]	1.1436	0.9301	1.4062
BMI	1.0001*	1.0000	1.0002
Model 1	HR	lower CI	upper CI
Weak[a]	3.948***	3.474	4.485

[a] Non-weak is reference [b] Male is reference [c] College and above is reference [d] White is reference

*** p<0.01 ** p < 0.05 * p < 0.1

After K-M method, I fit unadjusted and adjusted Cox models (Model 1 and Model 2) with covariates age, gender, education, race and BMI (Table 2 above). The unadjusted model had a hazard ratio 3.94 which is higher than the one in the study (1.59). The adjusted model 2 has muscle weakness HR of 2.02 which a significant. The model 1 HR means that muscle weakness is associated with the risk of death and having a weak increase the risk of death by 3.69 times at any particular time. Other estimated hazard ratios of covariates are very close to the one in the paper.

As mentioned in part 2, one of issues with the study was adding new participants into the study and assuming they have the same probability of survival with other participants who joined more than 4- 8 years ago. Another issue is with creating binary dependent muscle weakness variable (grip strength) from quantitative variable. This change makes the Cox model to ignore any change that happens above cut-points. Therefore, my model 1 will run Cox model for the first cohort which is between 2006 – 2010 with covariates age, gender, BMI, race, education and fixed grip-strength. The second model will fit Cox model for all the cohorts with time-varying grip-strength without categorizing it.

Table 3: New Model 1 and Model 2

	Model 1		**Model 2**	
	Hazard Ratio	*95% CI*	*Hazard Ratio*	*95% CI*
Weak[a]	0.99	0.98, 1.01	1	0.9999 , 1.0001
Age	1.08***	1.07, 1.09	1.10***	1.0957 , 1.1114
Female[b]	0.57***	0.46, 0.70	0.63***	0.5553 , 0.7196
BMI	0.99	0.98, 1.00	1.00	1.0000 , 1.0002
Race[d]	1.07	0.86, 1.34	1.36**	1.1132 , 1.6693
Education[c]				
GED	0.87	0.61, 1.26	1.81***	1.3778 , 2.3915
High School	0.74 **	0.62, 0.88	1.38***	1.1752 , 1.6267
Some College	0.48**	0.27, 0.83	1.13	0.7907 , 1.6289

[a] Non-weak is reference [b] Male is reference [c] College and above is reference [d] White is reference

*** p<0.01 ** p < 0.05 * p < 0.1

The first model shows that Hazard ratio for the handgrip is not significant. However, it shows that other covariates are significant (Table 3). I run the same model 1 with the depended variable (muscle strength) being categorial (weak vs. non-weak) and the output of the model is given in Appendix under new Model 1 v2. The result shows a significant hazard ratio estimator for the muscle weakness (variable weak).

The second model also shows that the handgrip measure is not significant. The handgrip strength becomes significant when we change from quantitative into categorial variable (weak and non-weak) and fits a new model. I think that changing the handgrip strength makes the biggest difference in the paper that is published and it should be examined more to check if truly there is a relationship between muscle strength and mortality of older people.

Reference

Goel, M. K., Khanna, P., & Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. International journal of Ayurveda research, 1(4), 274-8.

Kate Duchowny; Do Nationally Representative Cutpoints for Clinical Muscle Weakness Predict Mortality? Results From 9 Years of Follow-up in the Health and Retirement Study, The Journals of Gerontology: Series A, , gly169, https://doi.org/10.1093/gerona/gly169

Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., & Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery, 143(3), 331-6.

Time-To-Event Data Analysis, Columbia University 2018. Retrieved from https://www.mailman.columbia.edu/research/population-health-methods/time-event-data-analysis

Appendix

HHID	PN	TIME4	TIME5	handgrip	AGE
3	10	0	48	Non-Weak	82
3	10	48	95	Weak	82
3	20	0	48	Weak	80
3	20	48	95	Non-Weak	80
3	20	95	96	Weak	80
10004	10	0	48	Non-Weak	79
10004	10	48	69	Non-Weak	79
10004	40	0	48	Non-Weak	72
10004	40	48	95	Non-Weak	72
10004	40	95	96	Non-Weak	72
10013	10	0	48	Weak	80
10013	10	48	95	Weak	80
10013	40	48	95	Non-Weak	71
10013	40	95	96	Weak	71

Figure A1

Table A4

Table 1. Cutpoints for Clinical Muscle Weakness by Race and Sex in the Health and Retirement Study

	White Males (n = 3,279)	Black Males (n = 422)	White Women (n = 4,286)	Black Women (n = 738)
	Cutpoint (kg)	Cutpoint (kg)	Cutpoint (kg)	Cutpoint (kg)
Weak	<35	<40	<22	<31
Normal	≥35	≥40	≥22	≥31

New Model 1 V2.

```
> summary(fit6)
Call:
coxph(formula = surv.obj4 ~ GRIPST + AGE + factor(GENDER) + factor(RACE) +
    BMI + EDU, data = h06i_r_2)

  n= 6409, number of events= 714
   (395 observations deleted due to missingness)

                          coef exp(coef)  se(coef)      z Pr(>|z|)
GRIPST                0.809355  2.246459  0.099632  8.123 4.53e-16 ***
AGE                   0.066418  1.068673  0.004406 15.073  < 2e-16 ***
factor(GENDER)2      -0.541443  0.581908  0.076441 -7.083 1.41e-12 ***
```

```
factor(RACE)2        -0.112351  0.893731  0.115457 -0.973   0.33050
factor(RACE)7         0.432412  1.540970  0.229385  1.885   0.05942 .
BMI                  -0.004904  0.995108  0.005055 -0.970   0.33197
EDUCollege and above -0.667374  0.513054  0.124806 -5.347 8.93e-08 ***
EDUGED               -0.049643  0.951569  0.184330 -0.269   0.78769
EDUHigh School       -0.250698  0.778257  0.089217 -2.810   0.00495 **
EDUSome college      -0.630416  0.532370  0.277905 -2.268   0.02330 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                     exp(coef) exp(-coef) lower .95 upper .95
GRIPST                  2.2465     0.4451    1.8480    2.7309
AGE                     1.0687     0.9357    1.0595    1.0779
factor(GENDER)2         0.5819     1.7185    0.5009    0.6760
factor(RACE)2           0.8937     1.1189    0.7127    1.1207
factor(RACE)7           1.5410     0.6489    0.9830    2.4157
BMI                     0.9951     1.0049    0.9853    1.0050
EDUCollege and above    0.5131     1.9491    0.4017    0.6552
EDUGED                  0.9516     1.0509    0.6630    1.3657
EDUHigh School          0.7783     1.2849    0.6534    0.9270
EDUSome college         0.5324     1.8784    0.3088    0.9178

Concordance= 0.746  (se = 0.011 )
Rsquare= 0.089   (max possible= 0.856 )
Likelihood ratio test= 594.9  on 10 df,   p=<2e-16
Wald test            = 545.6  on 10 df,   p=<2e-16
Score (logrank) test = 605.5  on 10 df,   p=<2e-16
```

Table A5: Original adjusted model from the paper.

Table 3. Extended Cox Proportional Hazard Models: Association Between Muscle Weakness and Mortality in the Health and Retirement Study (*N* = 8,326), 2006–2014

	Model 1		Model 2		Model 3[f]	
	Hazard Ratio	95% CI	Hazard Ratio	95% CI	Hazard Ratio	95% CI
Weak[a]	1.59***	1.42, 1.77	1.52***	1.15, 1.47	1.32***	1.17, 1.48
Demographic factors						
Age (years)			1.02**	1.01, 1.03	1.00	0.99, 1.01
Female[b]			0.73***	0.66, 0.81	0.67***	0.60, 0.74
Black[c]			0.86	0.74, 1.01	0.87	0.74, 1.01
Education[d]						
GED			0.78*	0.60, 1.00	0.78	0.61, 1.02
High School			0.71***	0.63, 0.82	0.75***	0.65, 0.85
Some college			0.71***	0.63, 0.83	0.78**	0.67, 0.90
College and above			0.54***	0.46, 0.64	0.59***	0.50, 0.70
Smoking status[e]						
Former			1.42***	1.27, 1.59	1.32***	1.18, 1.49
Current			2.06***	1.70, 2.49	1.60***	1.31, 1.96