Trach/PCF Data Analysis

John Randazzo

8/17/2017

Methods

Univariate tests for association with tracheotomy timing were carried out using t-tests for normally distributed continuous variables, Wilcoxon Rank Sum Test for non-normally distributed continuous variables, Chi-squared tests for categorical variables and Fisher's Exact test for categorical variables when appropriate. Details are presented on the individuals who deceased in while under observation.

Logistic regression was carried out for mortality, pneumonia, acute respiratory failure, and sepsis, including tracheotomy timing as an independent variable and adjusting for risk factors that proved to be significant in predicting a particular complication. Models were selected via a backwards stepwise progression using R's step() function with the aid of MuMIn's dredge() function, using Akaike's Information Criterion (AIC) to assess the relative quality of different models. Kaplan-Meier plots for early and late tracheostomy groups are presented with Log rank p-values. These show vent days, hospital length of stay, and ICU length of stay as response variables with mortality as a censoring event. A Cox Proportional Hazards regression model for ICU length of stay is also presented, arrived at by the same means as the presented Logistic regression models, and associated tests for proportionality (see: survival's cox.zph() function) for predictors indicate the model's accordance with the assumption of proportional hazards for Cox regression.

Section 1: Data Overview/Summary Statistics

Summary of Numeric Data

Data were partitioned into two groups, subjects with Trach POD < 5 form the "early" group, and those with Trach POD $>= 5$ fall into the "late" group.

Table 1: Summary of Trach Timing

	0%	25%	50%	75%	100%
Early	0	1	2	3	4
Late	5	6	8	11	21

Table 2: Age

	0%	25%	50%	75%	100%
Early	15	35.5	59.0	78	92
Late	15	44.0	63.5	83	94

T-Test P-value = 0.4345108

Table 3: ISS

	0%	25%	50%	75%	100%
Early	5	16.5	25	36.5	75

	0%	25%	50%	75%	100%
Late	2	16.0	25	30.0	75

Wilcoxon Rank-Sum P-value = 0.3654124

Table 4: Head/Neck Score

	0%	25%	50%	75%	100%
Early	2	3	5	5	6
Late	1	3	4	5	6

Wilcoxon Rank Sum P-value = 0.0288141

Table 5: Face Score

	0%	25%	50%	75%	100%
Early	0	0	0	0	2
Late	0	0	0	0	3

Wilcoxon Rank-Sum P-value = 0.9784622

Table 6: Chest Score

	0%	25%	50%	75%	100%
Early	0	0	0	2	5
Late	0	0	0	3	5

Wilcoxon Rank-Sum P-value = 0.5210888

Table 7: Abdomen/Pelvic Score

	0%	25%	50%	75%	100%
Early	0	0	0	0	4
Late	0	0	0	0	5

Wilcoxon Rank-Sum P-value = 0.5466772

Table 8: Hospital LOS

	0%	25%	50%	75%	100%
Early	7	17	24	37.00	86
Late	8	23	29	37.75	112

Wilcoxon Rank-Sum P-value = 0.0066494

Table 9: ICU LOS

	0%	25%	50%	75%	100%
Early	0	10	16	22.00	71
Late	5	14	19	24.75	108

Wilcoxon Rank-Sum P-value = 0.0041306

Table 10: Vent Days

	0%	25%	50%	75%	100%
Early	0	9	18	26.25	73
Late	1	12	19	26.00	108

Wilcoxon Rank-Sum P-value = 0.5918962

Table 11: Admit GCS

	0%	25%	50%	75%	100%
Early	3	4	11	15	15
Late	3	9	15	15	15

Wilcoxon Rank-Sum P-value = 0.0397064

Summary of Categorical Data

Table 12: Obesity

	Early	Late	Early	Late
Not Obese	80	99	0.9638554	0.9705882
Obese	3	3	0.0361446	0.0294118

Fisher's Exact Test P-value = 1

Table 13: Pre-existing Diabetes

	Early	Late	Early	Late
No	65	87	0.7831325	0.8529412
Yes	18	15	0.2168675	0.1470588

Chi-Square Test P-value = 0.2981176

Table 14: Pre-existing Pulmonary Condition

	Early	Late	Early	Late
No	73	97	0.8795181	0.9509804

	Early	Late	Early	Late
Yes	10	5	0.1204819	0.0490196

Chi-Square Test P-value = 0.1335435

Table 15: Pre-existing Cardiac Condition

	Early	Late	Early	Late
No	49	47	0.5903614	0.4607843
Yes	34	55	0.4096386	0.5392157

Chi-Square Test P-value = 0.108176

Table 16: SCI Level

	Early	Late	Early	Late
High (C1-C4)	61	67	0.7349398	0.6568627
Low (C5-C7)	22	34	0.2650602	0.3333333
None	0	1	0.0000000	0.0098039

Chi-Square Test P-value = 0.3737691

Table 17: SCI Complete/Incomplete

	Early	Late	Early	Late
Complete	48	43	0.5783133	0.4215686
Incomplete	16	19	0.1927711	0.1862745
Intact	19	40	0.2289157	0.3921569

Fisher's Exact Test P-value = 0.0468961

Table 18: Discharge Destination

	Early	Late	Early	Late
Home	1	3	0.0138889	0.0322581
Long Term Acute Care Center	2	9	0.0277778	0.0967742
Other Hospital	3	1	0.0416667	0.0107527
Rehab Center	53	63	0.7361111	0.6774194
Skilled Nursing Facility	13	17	0.1805556	0.1827957

Fisher's Exact Test P-value = 0.2739727

Table 19: Mechanism of Injury

	Early	Late	Early	Late
Fall	47	60	0.5662651	0.5882353
MVA	32	40	0.3855422	0.3921569

	Early	Late	Early	Late
OBFT	4	2	0.0481928	0.0196078

Fisher's Exact Test P-value = 0.5989271

Table 20: Fusion Level (H/L)

	Early	Late	Early	Late
High	74	87	0.8915663	0.8613861
Low	9	14	0.1084337	0.1386139

Chi-Square Test P-value = 0.6950768

Table 21: Acute Respiratory Failure

	Early	Late	Early	Late
No ARF	65	58	0.7831325	0.5686275
ARF	18	44	0.2168675	0.4313725

Chi-Square Test P-value = 0.0035286

Table 22: Soft Tissue Infection

	Early	Late	Early	Late
No STI	82	99	0.9879518	0.9705882
STI	1	3	0.0120482	0.0294118

Fisher's Exact Test P-value = 0.6289327

Table 23: Infection

	Early	Late	Early	Late
No Infection	81	93	0.9759036	0.9117647
Infection	2	9	0.0240964	0.0882353

Fisher's Exact Test P-value = 0.1149122

Table 24: Sepsis

	Early	Late	Early	Late
No Sepsis	62	63	0.746988	0.6176471
Sepsis	21	39	0.253012	0.3823529

Chi-Square Test P-value = 0.0870444

Table 25: Mortality

	Early	Late	Early	Late
Alive	72	93	0.8674699	0.9117647
Dead	11	9	0.1325301	0.0882353

Chi-Square Test P-value = 0.4672528

Table 26: Pneumonia

	Early	Late	Early	Late
No Pneumonia	29	27	0.3493976	0.2647059
Pneumonia	54	75	0.6506024	0.7352941

Chi-Square Test P-value = 0.2774121

Table 27: ARDS

	Early	Late	Early	Late
No ARDS	79	96	0.9518072	0.9411765
ARDS	4	6	0.0481928	0.0588235

Chi-Square Test P-value = 1

Table 28: Perc or Open Trach

	Early	Late	Early	Late
Open	16	17	0.1927711	0.1666667
Perc	67	85	0.8072289	0.8333333

Chi-Square Test P-value = 0.7885395

Table 29: Initial Methylprednislone Dose

	Early	Late	Early	Late
No	54	70	0.6506024	0.6862745
Yes	29	32	0.3493976	0.3137255

Chi-Square Test P-value = 0.7217761

Table 30: Steroids

	Early	Late	Early	Late
No	44	59	0.5301205	0.5784314
Yes	39	43	0.4698795	0.4215686

Chi-Square Test P-value = 0.6106889

Details on the Subjects that Died Under Observation

Age	Early/Late	Cardiac?	Vent Days	ICU LOS	Hosp LOS	Admit GCS
85	Late	Yes	42	10	42	15
84	Early	Yes	32	19	39	8
78	Early	Yes	51	10	51	8
70	Late	No	54	54	54	3
86	Late	Yes	15	45	45	15
77	Early	No	9	9	9	3
77	Early	Yes	34	35	37	15
78	Late	Yes	19	19	19	5
89	Late	Yes	15	16	16	14
59	Early	No	7	7	7	3
92	Early	Yes	21	20	21	15
84	Late	Yes	19	19	19	14
84	Early	Yes	35	34	35	10
90	Early	No	20	23	27	15
87	Late	Yes	22	21	22	15
54	Early	No	42	35	42	3
77	Early	Yes	8	8	8	3
94	Late	No	19	20	20	3
73	Late	Yes	20	22	23	15
69	Early	Yes	10	11	11	15

Section 2: Logistic Regression

Presented are logistic regression models for 4 outcomes: Pneumonia, Sepsis, Mortality and Acute Respiratory Failure.

Table 32: Logistic Regression for Acute Respiratory Failure

	Est. Odds Ratio	95% CI Lower	95% CI Upper	P-Value
(Intercept)	0.3166345	0.1215682	0.7984305	0.0158862
groupLate	2.3909515	1.2205102	4.8081984	0.0123188
`Pre-existing Cardiac Condition`Yes	2.6161837	1.3187538	5.2727966	0.0063183
ISS	0.9792956	0.9563589	0.9994142	0.0589033

Table 33: Logistic Regression for Pneumonia

	Est. Odds Ratio	95% CI Lower	95% CI Upper	P-Value
(Intercept)	6.1143552	2.3147107	17.5762254	0.0004301
groupLate	1.4567609	0.7588273	2.8077212	0.2580771
Age	0.9731102	0.9550586	0.9903961	0.0030975
`Pre-existing Cardiac Condition`Yes	2.5848874	1.1376149	6.0559346	0.0252924

Table 34: Logistic Regression for Sepsis

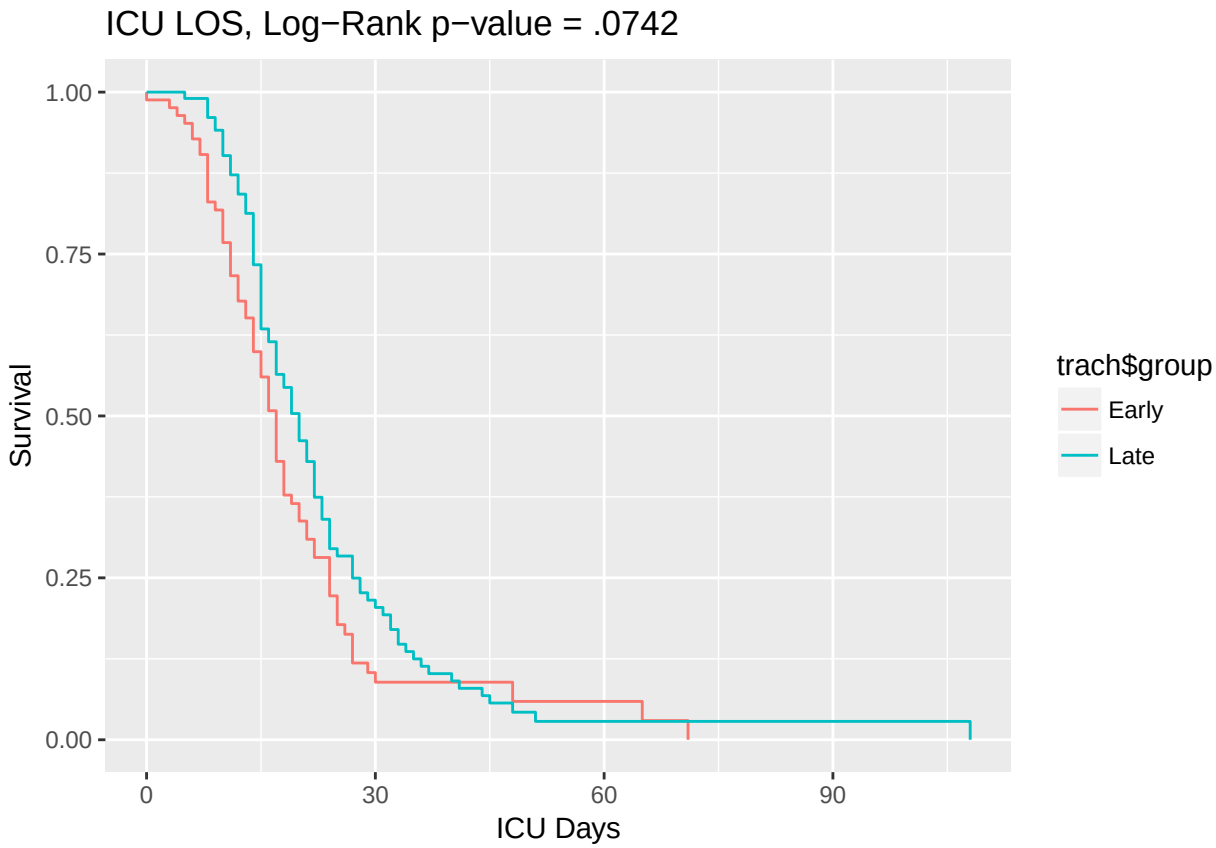
	Est. Odds Ratio	95% CI Lower	95% CI Upper	P-Value
(Intercept)	0.2396853	0.1274305	0.4260937	0.0000031
groupLate	1.8269382	0.9507025	3.5855256	0.0740200
`Pre-existing Cardiac Condition`Yes	1.6129130	0.7964921	3.2707494	0.1829906
`Pre-existing Diabetes`Yes	1.6796169	0.6996261	4.0436352	0.2439390

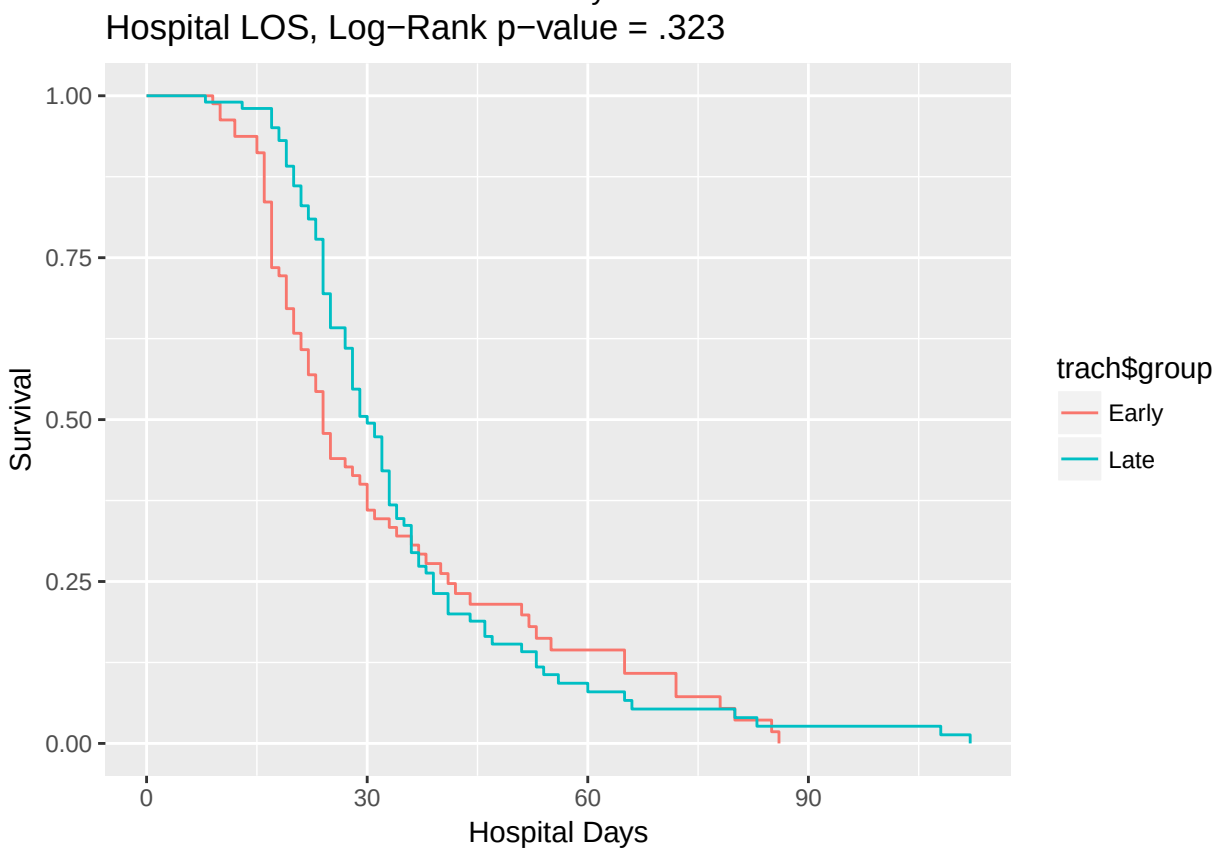
Table 35: Logistic Regression for Mortality

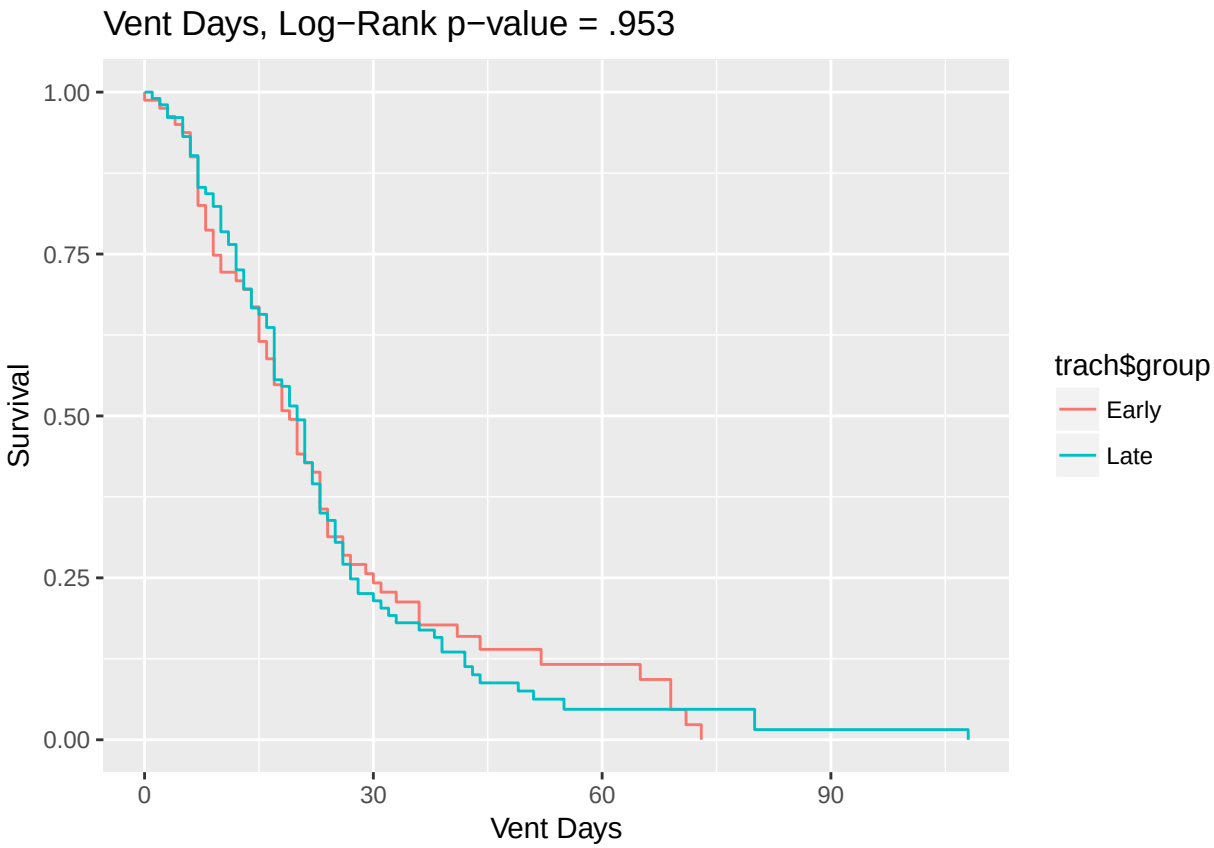
	Est. Odds Ratio	95% CI Lower	95% CI Upper	P-Value
(Intercept)	0.0001119	0.0000006	0.005578	0.0000909
groupLate	0.5782457	0.2023789	1.619241	0.2965015
Age	1.0968782	1.0494812	1.164676	0.0004386
ISS	1.0271034	0.9945951	1.059997	0.0910037

Results from Survival Analysis

Kaplan-Meier plots for the three time variables (Vent days, Hospital LOS, ICU LOS) are presented:


ICU LOS, Log−Rank p−value = .0742
Survival
1.00
0.75
0.50
0.25
0.00
0
30
60
90
ICU Days
trach$group
Early
Late


Hospital LOS, Log−Rank p−value = .323
Survival
1.00
0.75
0.50
0.25
0.00
0
30
60
90
Hospital Days
trach$group
Early
Late



First (non-Proportional Hazards) Cox Model

Presented are results from a Cox Proportional Hazards regression model for ICU LOS followed by results from proportionality checks amongst predictors:

	Est. Hazard Ratio	95%CI Lower	95%CI Upper	P-Value
groupLate	0.6833686	0.4938044	0.9457037	0.0216336
Age	0.9871060	0.9797853	0.9944814	0.0006331
`Pre-existing Pulmonary Condition`Yes	0.5009617	0.2683013	0.9353758	0.0300325
ISS	0.9895357	0.9793702	0.9998068	0.0458643
`Gender Code`Male	0.7226007	0.5063448	1.0312179	0.0733656

This model violates the Proportional Hazards assumption, evidenced by low P-values (in the following table) which correspond to the set of hypotheses:

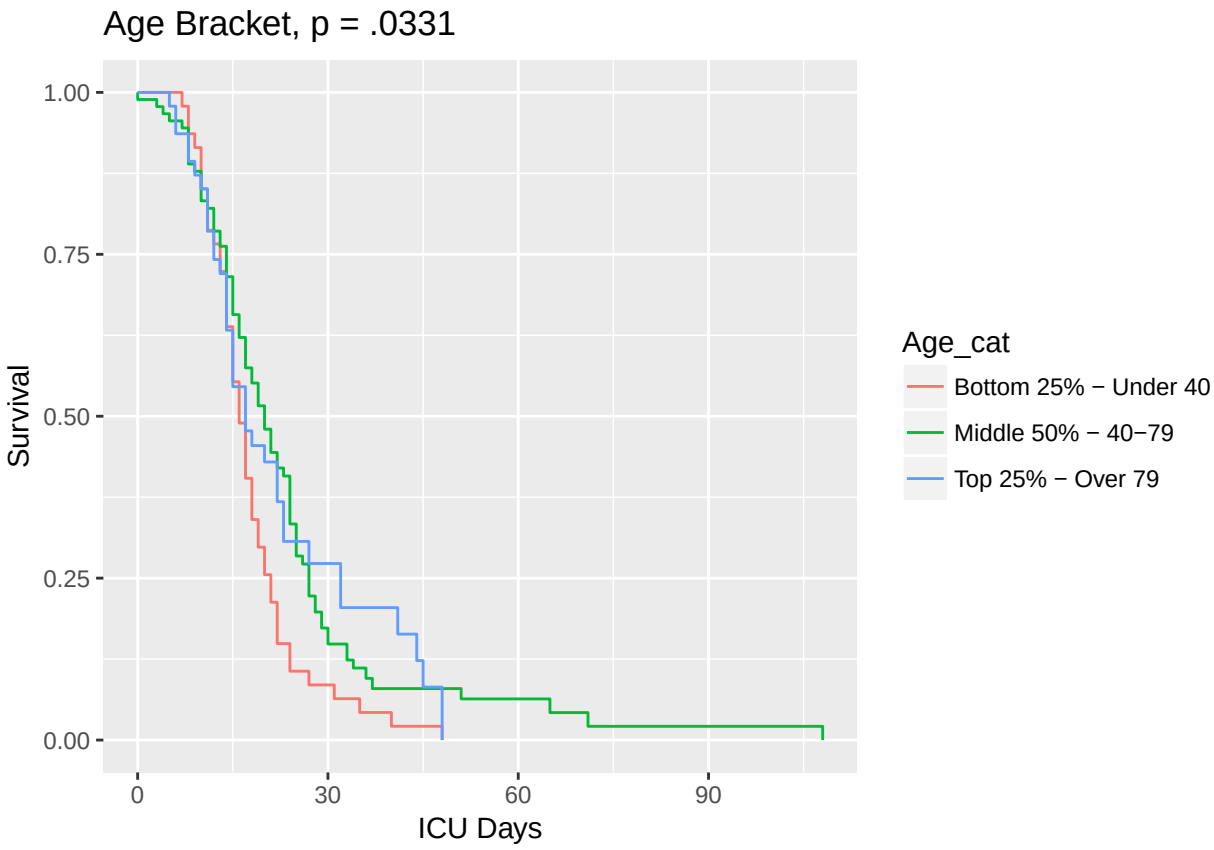
H_null = The variable meets the PH assumption
H_alternative = The variable does not meet the PH assumption

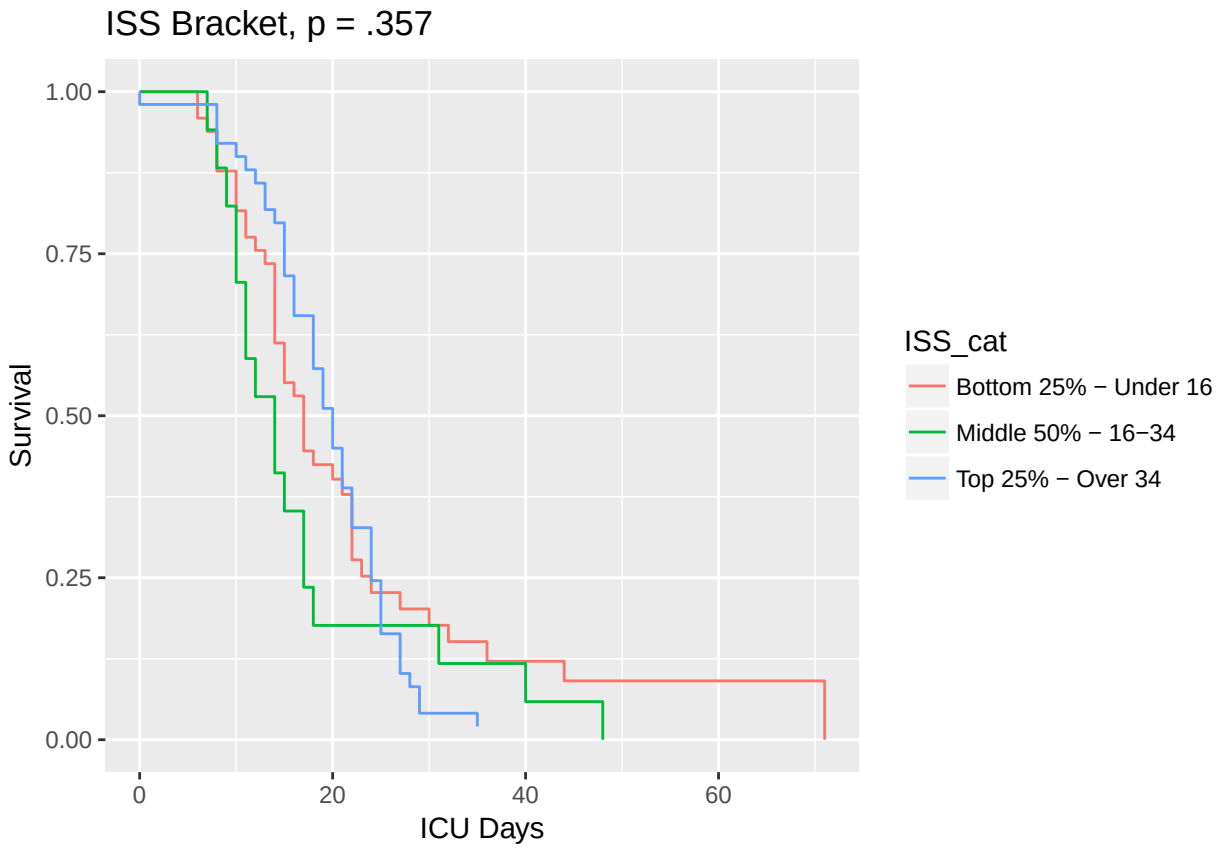
	Proportional Hazards P-Value
groupLate	0.1108462
Age	0.7840932
`Pre-existing Pulmonary Condition`Yes	0.1211004
ISS	0.0013458
`Gender Code`Male	0.9292058

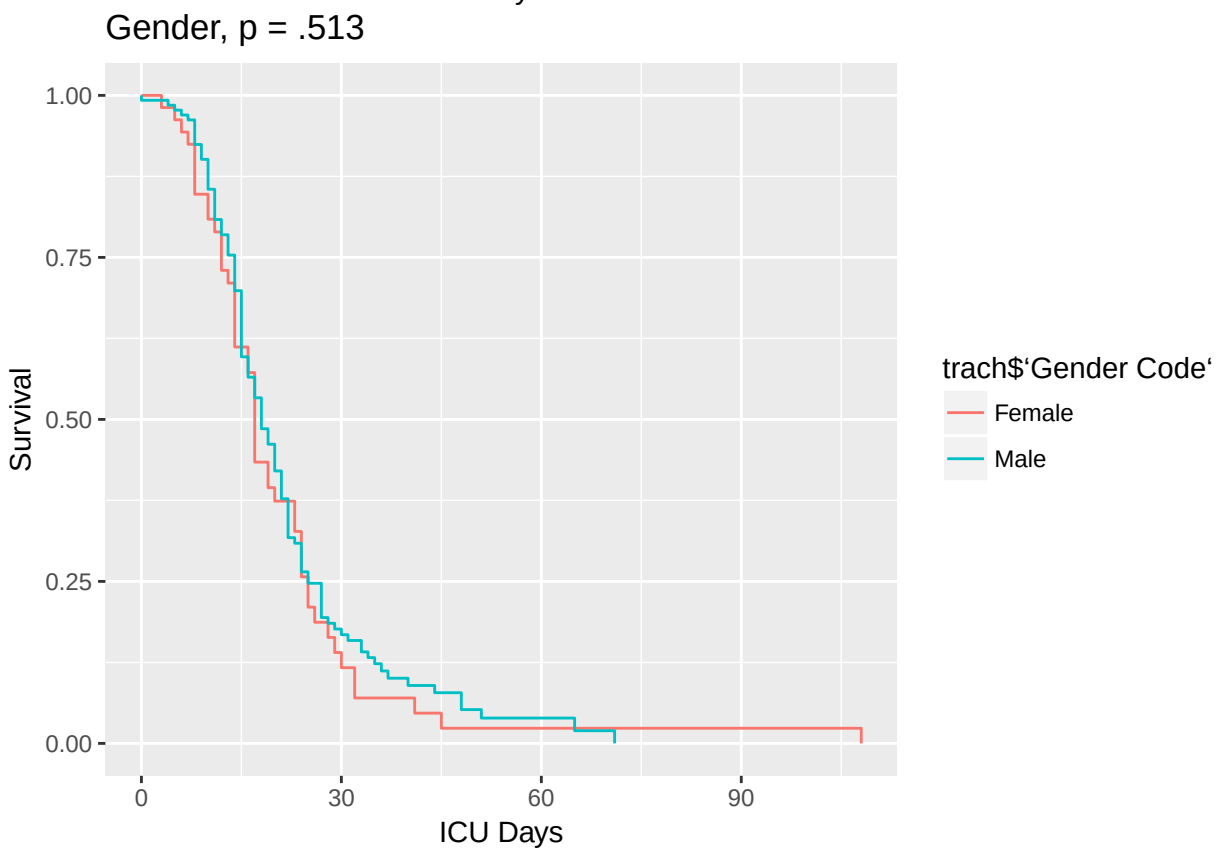
	Proportional Hazards P-Value
GLOBAL	0.0027746

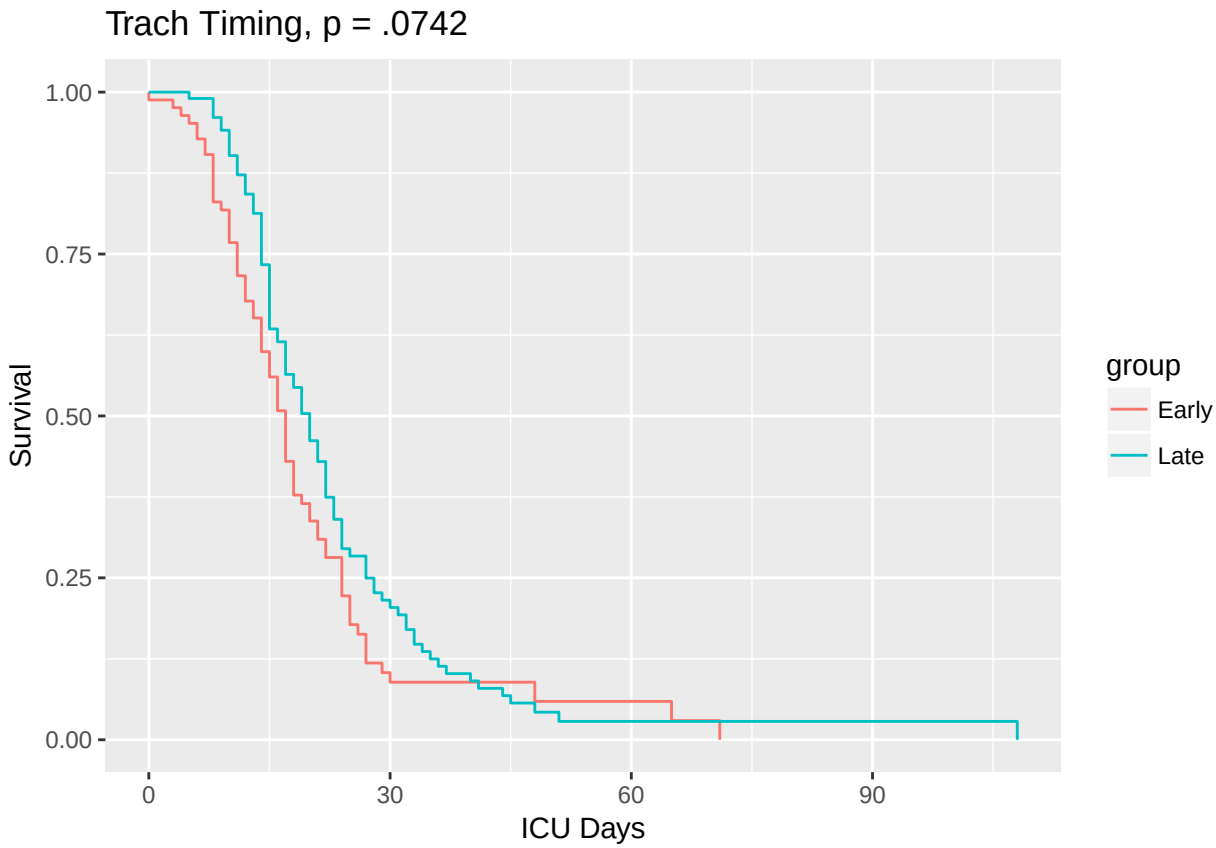
Each variable as well as the entire model are tested, and there is compelling evidence that ISS cannot be used in the traditional sense of Cox regression, and so it was removed from consideration for model building. (There are methods to include this variable as a predictor if desired.)

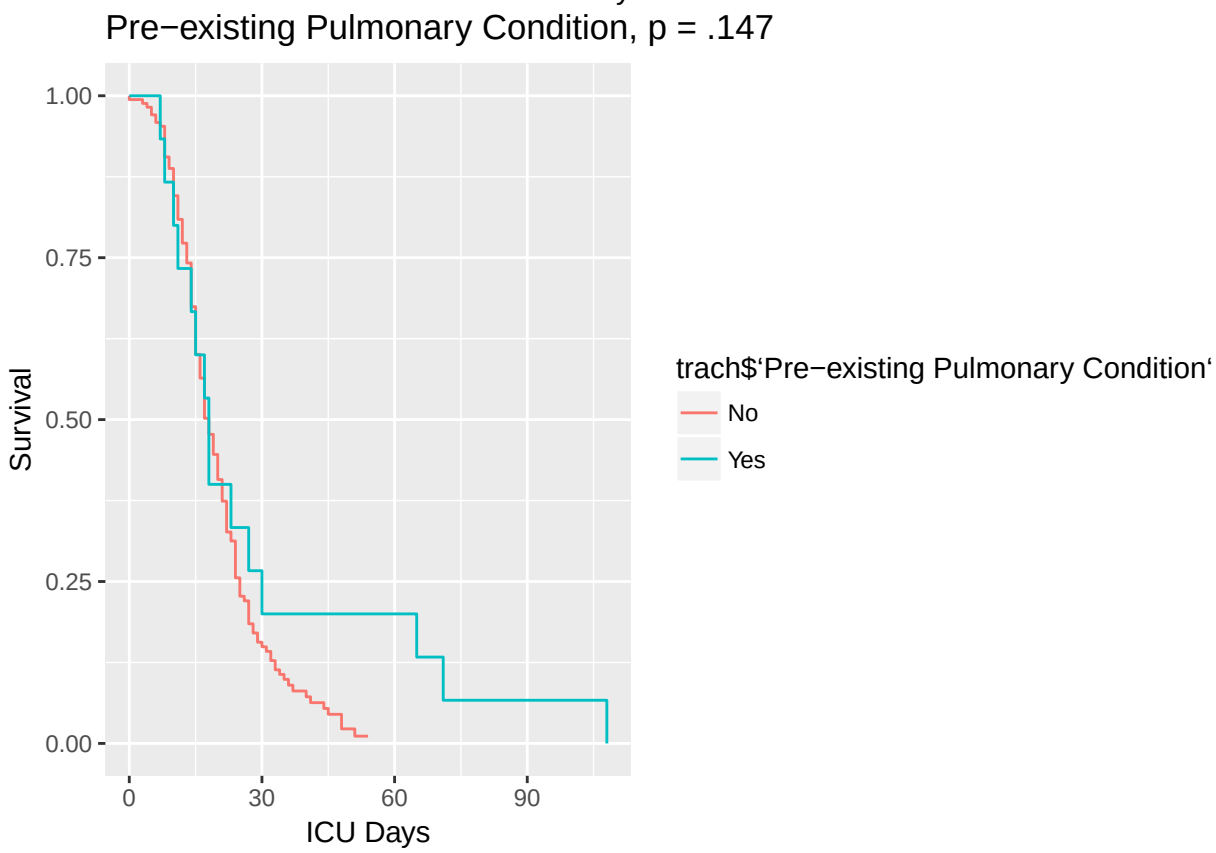
Presented are Kaplan Meier estimates which compute separate survival curves for different levels of each unique predictor variable included in this model, along with their associated p-values from the log-rank test:




ISS Bracket, p = .357
1.00
0.75
0.50
0.25
0.00
Survival
0
20
40
60
ICU Days
ISS_cat
Bottom 25% − Under 16
Middle 50% − 16−34
Top 25% − Over 34


Gender, p = .513
1.00
0.75
0.50
0.25
0.00
Survival
0
30
60
90
ICU Days
trach$‘Gender Code‘
Female
Male





Note the skewed tail end for the group with pre-existing pulmonary condition; all of the “outlier” values

(extremely long stints in ICU) have the pre-existing pulmonary condition. It is likely that this variable is only significant in our Cox model due to these extreme values.

Details on the Subjects with ICU LOS > 60

Table 38: Details on Subjects With ICU LOS > 60

ICU Days	Early/Late?	Age	Gender	Pulmonary?	ISS
108	Late	60	Female	Yes	27
65	Early	63	Male	Yes	25
71	Early	76	Male	Yes	10

Examination of the subjects with the 3 longest ICU stints (ICU LOS > 60) revealed that each of them had a pre-existing pulmonary condition. These 3 observations confounded the first model since their "survival time" carried a great deal of weight.

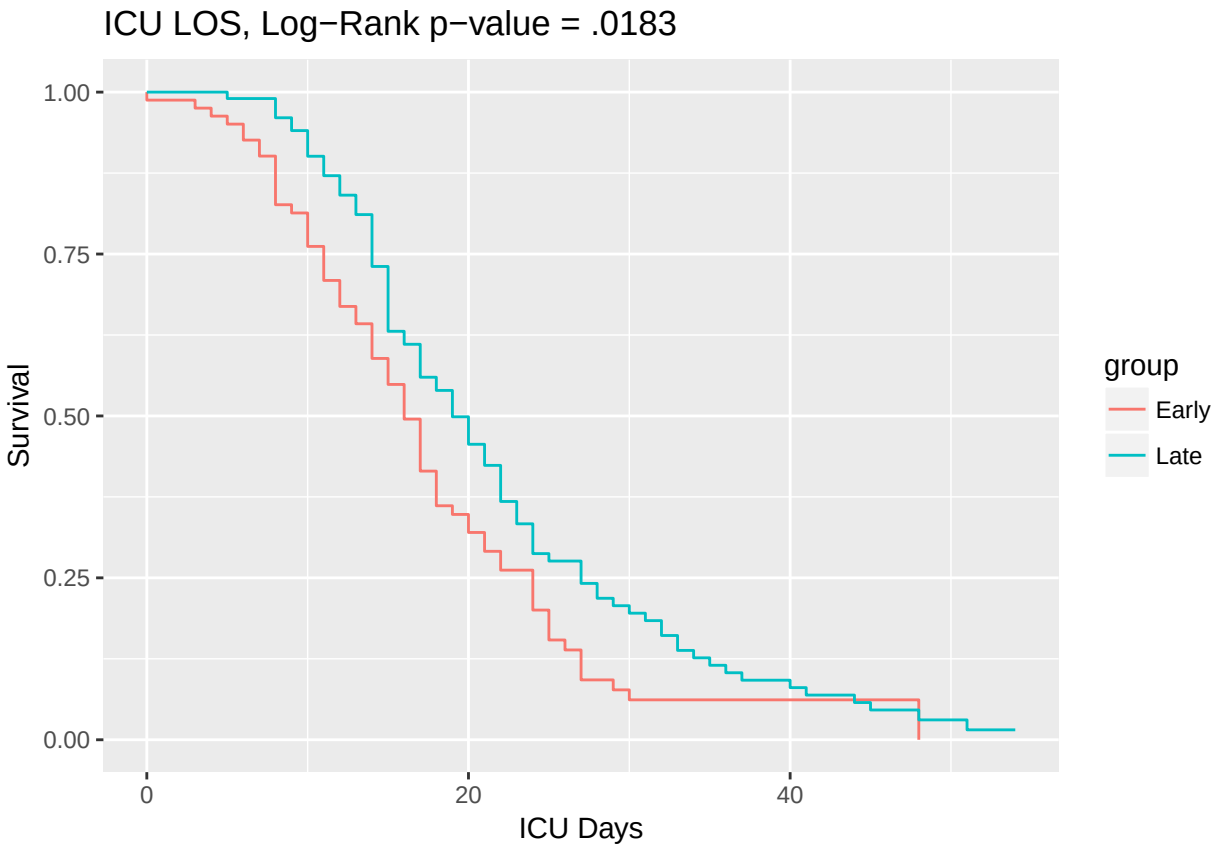
The Second Proportional Hazards Model

There are a number of ways to adjust a Cox model to meet the assumption, the most basic of which is removing extreme time values (i.e., extremely long ICU stints.) As stated in the previous section, there was a noticeable association with extreme ICU LOS values for individuals with a pre-existing pulmonary condition. Follwing the removal of the 3 subjects with extremely long ICU stints, the dataset shrinks from 185 to 182 total observations, and we re-examine the distribution of ICU LOS across the early and late trach groups, as well as the Kaplan-Meier estimate:

Table 39: ICU LOS Summary After Removing Subjects With ICU LOS > 60

	0%	25%	50%	75%	100%
Early	0	10	16	21	48
Late	5	14	19	24	54

Wilcoxon Rank-Sum P-value = 0.0018703



The following model was then obtained:

	Est. Hazard Ratio	95%CI Lower	95%CI Upper	P-Value
groupLate	0.7178860	0.5215425	0.9881464	0.0420415
Age	0.9908787	0.9841722	0.9976309	0.0081808
`Gender Code`Male	0.7387083	0.5199199	1.0495653	0.0910260

I will do my best to explain these results:

If there are two subjects that have the exact same measures (same age and gender) but differ in the variable of interest (subject A has early trach, subject B has late trach), then the hazard ratio (.7178) means that the late trach subject is estimated to be 71.78% as likely to experience the "event" of being released from the ICU compared to a subject with early trach, so early trach generally implies a shorter ICU stay. It should be noted that the late trach subject is estimated to be 71.78% as likely to leave the ICU at ANY time, as hazard rates are interpreted to be "instantaneous."

Similarly, with all other variables held equal, if subject B is one year older than subject A, then at any time during observation, subject B is estimated to be 99.08% as likely to experience the event of being released from the ICU as subject A. So old age implies a longer ICU stay.

A male subject with the same age and trach timing as a female subject in the study is estimated to be 73.87% as likely to leave the ICU at any time compared to the female counterpart.

It should be noted that the variable for pre-existing pulmonary condition was used in consideration for this model, but was shown to not be particularly useful.

Assessment of the model's alignment with the proportional hazards assumption indicated no clear violation:

	Proportional Hazards P-Value
groupLate	0.1566089
Age	0.1722266
`Gender Code`Male	0.6073447
GLOBAL	0.3233449

More on the proportional hazards assumption from http://www.phusewiki.org/docs/Conference%202013%20SP%20Papers/SP07.pdf: "The assumption states that hazard ratio for two subjects who are characterized by different sets of covariates depends only on values of these covariates and does not depend on time. In other words: hazard ratio is constant over time which means that the effect of the given covariate on the hazard level is the same at all timepoints. There are various opinions on the importance of this assumption with regard to the parameters interpretation. Some authors state that violation from it is nothing extremely problematic as in such cases parameter for a covariate for which assumption is not satisfied can be understood as 'average effect' over timepoints that are observed in a dataset (Allison, 1995). The others however underline the importance of this assumption (Hosmer, Lemeshow 1999) and suggest potential modification of the model if hazard ratio turns out not to be constant over time for some covariates."

Thank you for reading. If anything is unclear please email me at johnrandazzo1996@gmail.com.