Project 3: Death Rates across 44 VA Hospitals

Rachel Weber

November 24, 2018

Background

There are 26,489 patients in the data set, spanning 6 consecutive six month periods. The data contain patient information for height, weight, BMI, albumin levels, ASA, procedure code, and whether the patient died within 30 days. Table 1 below summarizes patient info, stratified by death within 30 days.

Table 1: Patient Information by Death within 30 Days

	Survived N = 25,259	**Died N = 693**
	N(%)	*N(%)*
Procedure		
0	4,951(20)	102(15)
1	20,308(80)	591(85)
Missing	0(0)	0(0)
ASA		
2	1,130(4)	13(2)
3	7,718(31)	102(15)
4	14,473(57)	485(70)
5	1,301(5)	66(10)
Missing	637(3)	27(4)
	Mean(SD)	*Mean(SD)*
Weight (lbs.)	165.8(36.03)	174.3(38.24)
Missing (N%)	620(2)	40(6)
Height	65.33(2.625)	65.51(2.575)
Missing (N%)	620(2)	40(6)
BMI	28.2(4.010)	29.49(4.058)
Missing (N%)	622(2)	40(6)
Albumin	4.041(0.5798)	4.031(0.5574)
Missing (N%)	12,623(50)	343(49)

As seen above, albumin had a large amount of missing data. Knowing this was a lab result typically not included until recently, some missingness is expected. To ensure that the missingness is random, and will not impact our findings, we inspected the data for patterns in missing values. Figure 1 in the

appendix is one plot of many used to check for trends. It shows, using box plots, the distribution of BMI for patients with (blue) and without (red) albumin measures. The boxplots are essentially identical, telling us that albumin missingness is not related to BMI. We also compared the percent of patients who died within 30 days. 6% of people missing albumin died within 30 days. This was highly similar to the percent of people with albumin measurements who died—4%. After much deliberation, it was decided that albumin is missing at random, and we need not be concerned about bias in our results due to its missingness.

Methods

Our objective was to determine if the death rates in VA hospitals during period 39 deviated from those we expected given historic data. Before analysis could begin, some data cleaning was necessary. BMI measures below 60 were considered by the investigator to be medically feasible. Any values above 60 were removed. It was also found that very few patients had an ASA of 1, so those with ASA 1 or 2 were combined into category 2. This was deemed medically reasonable given the highly similar status of patients in the two groups. A few patients had a procedure code of 2. We determined these instances were typos and those with such had their procedure codes removed. In some hospitals, weight was entered in kilograms instead of pounds. Since BMI for patients with such was still calculated using pounds, we were able to ignore these administrative errors and use BMI for modeling without adjustment. Since there are likely differences in the patient populations between hospitals, the mortality rates are risk-adjusted by hospital.

Modeling was performed using R version 1.1.453. A model predicting death within 30 days was built using logistic regression. Data to build the model contained only patients in periods prior to 39-- totaling 22,179 patients in the subset. Model covariates were BMI, Procedure, ASA, and Albumin. Later we did run a model excluding albumin, for comparison's sake. Non-parametric bootstrapping was used to repeatedly draw from the data (periods 34-38) and reconstruct the model with each sample. Samples

were ensured to be balanced for each hospital by drawing randomly from each hospital, with replacement, and then aggregating hospital data together. The model that resulted was used to predict the death rate within 30 days for time period 39. This was done for each hospital individually so risk-adjustment remained specific to a hospital's population. The mean of the predicted values for each hospital was aggregated over all bootstraps—1000 in total—and the mean of these means was taken to find the point estimate for predicted death rate in period 39. In addition, the 5^{th} and 95^{th} percentile of the means was collected to provide a reasonable estimate of variation for expected values.

Following the bootstrapped generation of a confidence interval, we used the binomial theorem to calculate the 95% confidence interval for death rate in period 39. Recalling that the variance of a binomial random variable is $p(1-p)$ with p being the observed death rate for a given hospital, we used the formula $CI = \mu \pm 1.96(SE) = \mu \pm 1.96\left(\sqrt{\frac{p(1-p)}{n}}\right)$ to calculate our confidence interval using observed death rates from just this final period. These two confidence intervals—that found from bootstrapping and that from the binomial theorem—were compared first to the observed rates for each hospital and then to each other for amount of overlap.

To assess how missingness in the albumin covariate was impacting expected rate estimates, a second bootstrap was run without albumin in the model. This bootstrap too was repeated 1000 times with proportional sampling from each hospital ensured.

Results

Table 2 in the appendix shows the confidence intervals and the observed death rates. Overall, the bootstrapped confidence intervals from the model containing albumin captured 32 of the 44 hospital's observed death rates for period 39. That is, 73% of the hospitals had observed rates within the expected variation seen from the predictive model. Of the 27% of hospitals that had values outside of the predicted rates, 9 had no patients die in period 39. The remaining 9% of hospitals had more deaths than

were expected given historic rates. They are shown in Figure 2 in the appendix. As a visual representation of Table 2, Figure 2 shows where the observed confidence intervals (blue) overlap the expected confidence intervals (red). Each black point is the observed death rate for each hospital—along the x-axis—in period 39. Naturally, the observed value will fall exactly in the middle of the observed confidence interval. What is relevant to see is where the observed does not overlap the expected. In these circumstances, it is reasonable to conclude that the hospital had an atypical month. For hospitals with no deaths in period 39, there is no blue observed interval to create, so while they deviate significantly from expected values, their deviation is good news. Hospitals 17, 21, and 34 have observed intervals that never or hardly overlap expected intervals. Hospital 34 had the worst deviation with 11.8% death rate in period 39 but a maximum expected death rate of 4.8% (p-value: <.0001). Hospitals 17 and 21 had slight overlap between expected and observed rates but still showed markedly high death rates (11.5 and 9.5, respectively). Though these are the most blatant deviations from expected values, p-values calculated by determining probability of observed number of deaths given expected death rate indicate that hospitals 7 and 14 also deviated significantly from that which was expected (table 2, highlighted). These are significant by the standard .05 threshold with p-values of 0.04 and 0.02, respectively.

Figure 3 shows the expected confidence intervals generated using a model that excluded albumin. Since patients dropped from the previous model due to missingness were included in this ancillary analysis, expected confidence intervals became significantly smaller. The precision increased dramatically with sample size, making it seem that most hospitals in period 39 deviated significantly from historic rates. For this reason, the model containing albumin was determined to be superior, despite the high missingness of the albumin variable.

Discussion

Despite high missingness in measures for albumin, modeling to predict hospital death rates given historic data produced reasonably tight confidence intervals when albumin was included in the model. Of the 44 hospitals studied, 32 had observed death rates within the expected range estimated from bootstrap sampling. Since single observations on each hospital gave no estimate of variability, the binomial theorem was employed to ensure that confidence in reasonable deviation from expected values remained high. Nine hospitals had zero deaths in period 39 and should be lauded for such. Three hospitals appeared to have significant deviations from those expected given historic data, and their observed confidence intervals failed to overlap with expected ranges.

You will note that the expected 95% CI for hospital 30 is NA in Table 2. This is likely because there were just 8 people who died for that hospital in all time periods prior to 39. There was almost no variation in this outcome, so it is likely the model failed to converge. Though we could not get an estimate for hospital 30, the extremely low death rate in period 39 (1.2%) fits the historically low rates we observed.

It is our recommendation that hospitals 7, 14, 17, 21, and 34 be contacted to determine why their death rates were higher than expected for this most recent period. In future death rate assessments, one should still investigate the data to ensure no variable is missing in concert with the value of another variable (or the outcome), though here no pattern exists. All code for analysis and figures can be found on the author's GitHub page.

Appendix

Figure 1. Distribution of BMI for all patients (far left), patients with albumin measures (blue) and patients missing albumin (red)

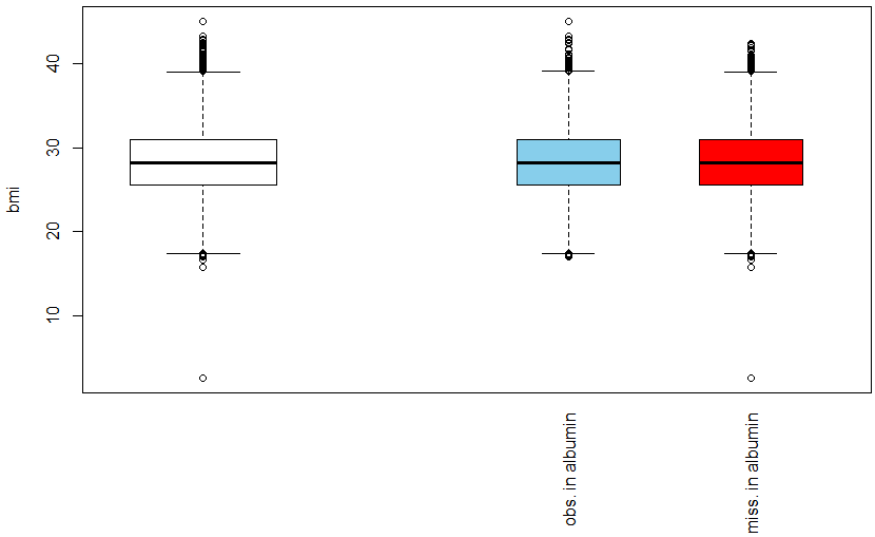

Table 2. Observed death rate (period 39) and the two 95% Cis generated in analysis.

Hospital	Obs. Death Rate (%)	95% Bootstrap CI	95% Binomial CI	P-value
1	1.89	0.59-5.08	0-4.48	0.2164
2	3.74	0.63-5	0.14-7.33	0.1028
3	2.2	0-4.83	0-5.21	0.2599
4	2.35	0.96-4.45	0-5.58	0.2484
5	3.06	0.6-4.74	0-6.47	0.2273
6	2.5	0.57-4.77	0-5.92	0.2706
7	6	0.81-4.42	1.35-10.65	0.0362
8	2.63	0.42-5.27	0-5.57	0.22
9	0	0.35-4.42	0-0	0.0817
10	2.65	0.51-4.75	0-5.62	0.2212
11	0.88	0.2-4.84	0-2.61	0.1411
12	1.96	0.84-4.63	0-4.65	0.2334
13	0.97	0.47-4.36	0-2.86	0.1339
14	7.59	0.78-4.82	1.75-13.44	0.0214
15	2.3	0.48-3.87	0-5.45	0.2717
16	3.48	0.74-5.55	0.13-6.83	0.1776
17	11.49	0.4-5.07	4.79-18.2	< 0.0001
18	0	0.2-4.86	0-0	0.1476
19	0	0.57-4.77	0-0	0.0808

20	1.08	0.97-4.92	0-3.17	0.1975
21	9.52	0.94-5.32	3.91-15.14	0.0013
22	2.97	0.34-5.6	0-6.28	0.2047
23	5.22	0.09-4.13	1.15-9.28	0.0647
24	1.01	0.54-4.39	0-2.98	0.1584
25	0	0.17-4.48	0-0	0.1089
26	1.75	0.05-5.22	0-4.16	0.1749
27	0	0.58-4.99	0-0	0.0767
28	3.77	0.36-4.83	0.15-7.4	0.1523
29	3.66	0.58-3.88	0-7.72	0.1966
30	1.16	NA	0-3.43	*NA*
31	0	0.87-4.27	0-0	0.0759
32	0	0.41-4.37	0-0	0.0846
33	1	0-5.2	0-2.95	0.2113
34	11.88	0.13-4.83	5.57-18.19	< 0.0001
35	2.25	0.26-5.55	0-5.33	0.2069
36	0	0-4.57	0-0	0.1382
37	1.92	0-5.16	0-4.56	0.272
38	1.06	0.3-4.71	0-3.14	0.2528
39	4.04	0.34-5.53	0.16-7.92	0.1607
40	0.92	0-4.46	0-2.71	0.2125
41	2.25	0.79-4.7	0-5.33	0.2711
42	1.82	0-5.28	0-4.32	0.2314
43	0	0.61-4.61	0-0	0.0915
44	0.85	0.39-4.4	0-2.52	0.2012

Figure 2. Confidence Intervals and Observed death rate. Bootstrapped intervals in red, binomial theorem intervals in blue

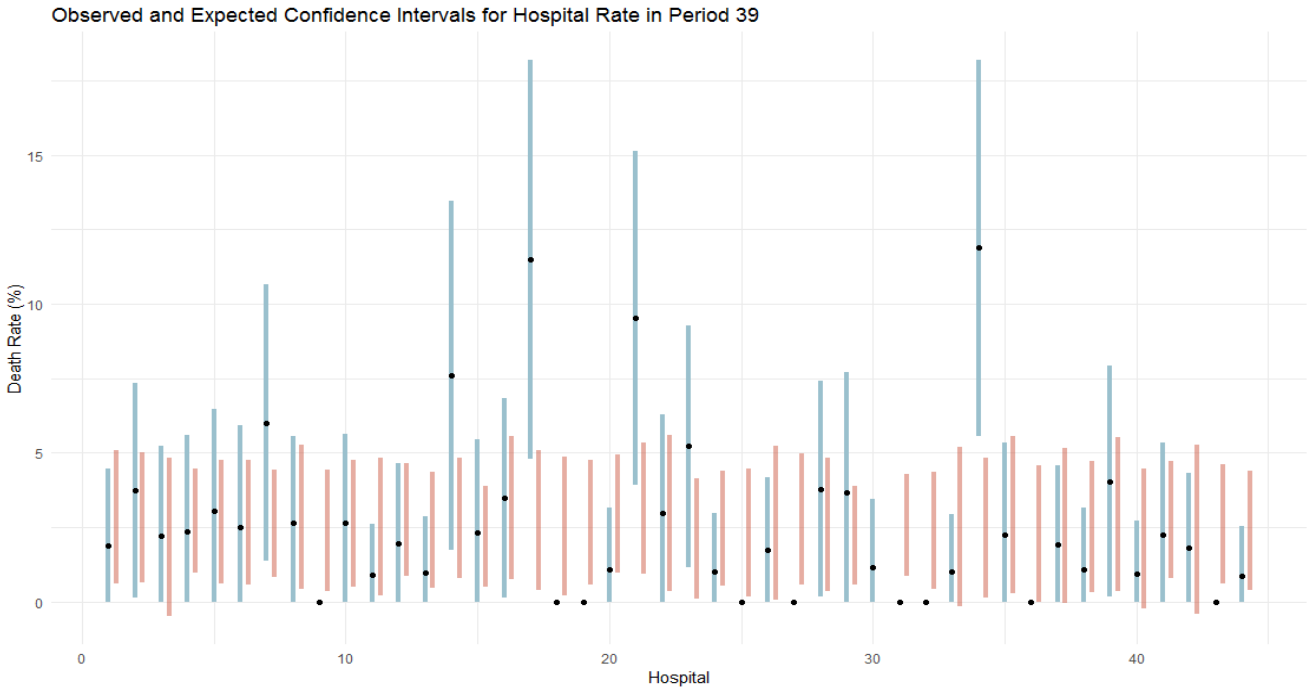

Figure 3. Confidence Intervals and Observed death rate. Bootstrapped intervals in red, binomial theorem intervals in blue. Bootstrapped intervals generated using a model without albumin