Smart Pill Executive Summary

Andrew Walther

12/7/2018

The Smart Pill Study

Introduction

Though delayed gastric emptying is a common problem among critically ill patients, it has proven difficult to monitor gastrointestinal function among mechanically ventilated (e.g. intubated) patients. This study, from Rauch et al., seeks to describe the first use of a new motility capsule as a means of monitoring gastric emptying, small bowel transit, and whole-gut transit times among the critically ill. The researchers used the capsule to compare these times for critically ill patients with intracranial hemorrhage to those times for healthy volunteers, as measured by the capsule.

To find their sample of critically ill patients, the researchers screened a hospital ICU twice a day. To be eligible, a patient had to be "intubated, ventilated, and [have] an Acute Physiology, Age, Chronic Health Evaluation score of greater than 25" (Rauch, et al., 2012). Patients under age 18, with IBD or abdominal trauma, or evidence of ileus or suspected obstruction were excluded, as were with pacemakers. Among the eligible patients, 8 were recruited within 36 hours of being admitted to the ICU. The control group was composed of 87 healthy volunteers with no history of major abdominal surgery.

The capsule was placed into the esophagus using a laryngoscope, and then advanced blindly into the stomach, at which point its position was radiographically confirmed. The healthy volunteers were studied in a separate, multicenter trial. Total GI transit time was calculated from the time the capsule was placed until it passed from the body. MotiliGI software was used the calculate gastric emptying. Small bowel transit time was calculated from the time the capsule entered the duodenum to the time it reached the cecum, with these changes in environment indicated by changes in pH and pressure (as measured by the capsule). Total GI transit time was calculated from the time the capsule was placed until it passed from the body.

Variables

The researchers gathered data on 22 variables, shown and classified below.

```
##    Variable Name          Classification
## 1  Group                  Categorical
## 2  Gender                 Categorical
## 3  Race                   Categorical
## 4  Height                 Numerical
## 5  Weight                 Numerical
## 6  Age                    Numerical
## 7  GE Time                Numerical
## 8  SB Time                Numerical
## 9  C Time                 Numerical
## 10 WG Time                Numerical
## 11 S Contractions         Numerical
## 12 S Sum of Amplitudes    Numerical
## 13 S Mean Peak Amplitude  Numerical
## 14 S Mean pH              Numerical
## 15 SB Contractions        Numerical
## 16 SB Sum of Amplitudes   Numerical
```

```
## 17 SB Mean Peak Amplitude Numerical
## 18 SB Mean pH             Numerical
## 19 C Contractions         Numerical
## 20 C Sum of Amplitudes    Numerical
## 21 C Mean Peak Altidue    Numerical
## 22 C Mean pH              Numerical
```

The variables of greatest interest to us were the primary response variables, and the explanatory variables that will help us understand them better. The primary response variables in this study are Gastric Emptying (GE) time, the time from ingestion to gastric emptying, and Small Bowel (SB) transit time, the time from gastric emptying to ileocecal junction. We also wanted to investigate the following explanatory variables further:

1) Stomach (S) Mean pH
2) Small Bowel (SB) Mean pH
3) Whole Gut (WG) time
4) Age.

Statistical Analysis

In analyzing the data, the demographics were summarized using means and standard deviations. They compared these sample statistics using the Student t or Fisher exact test. In this study, the Student t test was used on numerical response data, while the Fisher Exact test was used on the categorical data.

The primary outcomes of gastric emptying time and small bowel transit time were compared between critically ill patient and the healthy volunteers with Wilcoxon rank sum test.

With this data, the researchers were able to estimate a 97.5% confidence interval for the corresponding median differences. With only 8 critically ill patients and 87 controls, the study had approximately 85% power at the 0.025 significance level in order to detect a difference of more than 2.5 hours in gastric emptying and small bowel transit times (also assuming a standard deviation of 2 hours).

Paper Analysis

The researchers concluded that gastric emptying and small bowel transit times were markedly prolonged in critically ill trauma patients with intracranial hemorrhage, compared to normal volunteers. They therefore suggest that physicians caring for critically ill patients should consider a small intestinal dysfunction, along with delayed gastric emptying. Due to the variety of tests employed, evaluating both numerical and categorical data, as well as utilizing the Wilcoxon rank sum test in order to compare drastically different data sets, the researchers' conclusions appear justified.

Although the paper was generally succinct and gave explanations where needed without adding too much extraneous detail, two key details that would make the experiment more understandable to readers were left out. First, it is made clear that ICU patients were only accepted for the study after passing a "chronic health evaluation," but no description of that evaluation were given. Transparency on this evaluation method would ensure that the study is properly accepting or rejecting test subjects. Second, the researchers indicated that the demographic characteristics of the test subject group were well balanced and even provided a table with sex, age, height, and other characteristics, but there was no indication of how the researchers determined that the demographic characteristics were indeed "balanced." Note also that the results of this study were limited by its small sample size of critically ill patients.

Hypotheses Concerning Relevant Variables

Recall that the paper's primary response variables were Gastric Emptying (GE) time, the time from ingestion to gastric emptying, and Small Bowel (SB) transit time, the time from gastric emptying to the ileocecal junction.

We proposed hypotheses for each of the explanatory variables of interest mentioned above and their effect on the reponse variables.

> We propose that a decrease in Stomach Mean pH will lead to a decrease in GE time. A decrease in stomach pH should make the stomach a harsher chemical environment, leading to a more rapid breakdown of stomach contents and an increased rate for the early parts of the digestion process, including gastric emptying.
>
> We propose that a decrease in Small Bowel pH will lead to a decrease in SB time, for similar reasons.
>
> We propose that an increase in Whole Gut time will lead to an increase in both GE time and SB time.
>
> We propose that older patients (increased age) will have longer GE times and SB times, due to lower overall levels of digestive activity.

In addition to testing these four hypotheses, we also attempted to replicate the findings of the study using our own methods.

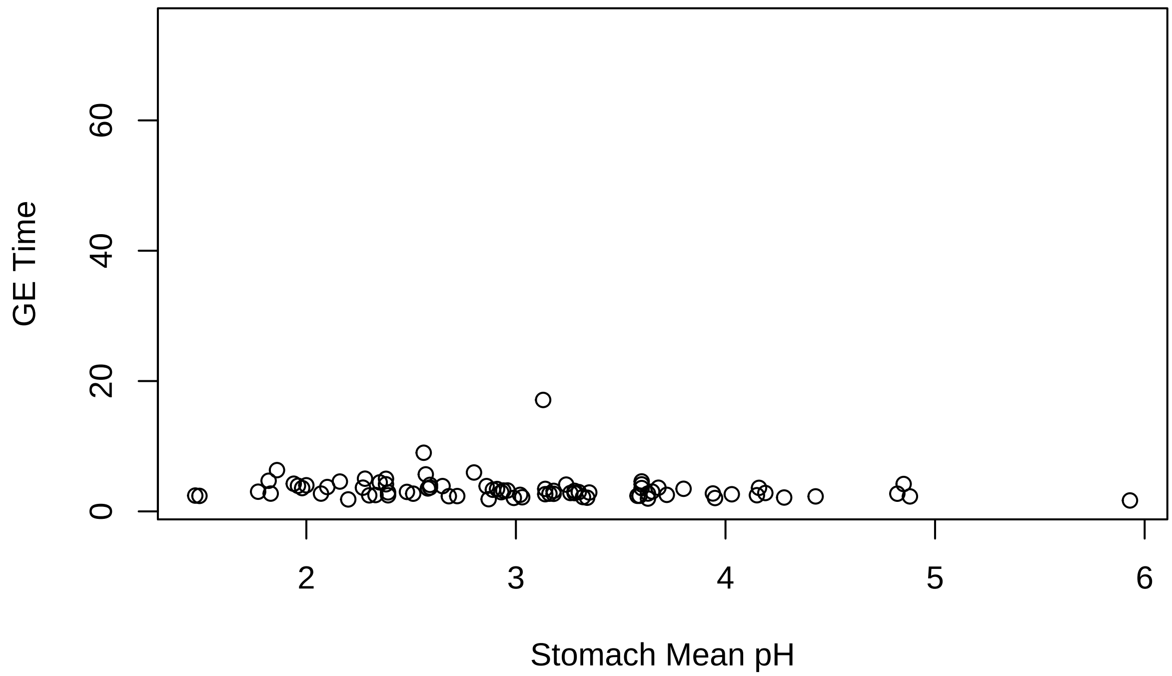
Variable Testing

All of our explanatory variables of interest were numerical, as were both response variables. Any hypothesis test for one of the eight possible relationships between these response and explanatory variables was performed using an appropriate technique such as correlation, linear regression, or multiple regression.

Stomach Mean pH and GE Time

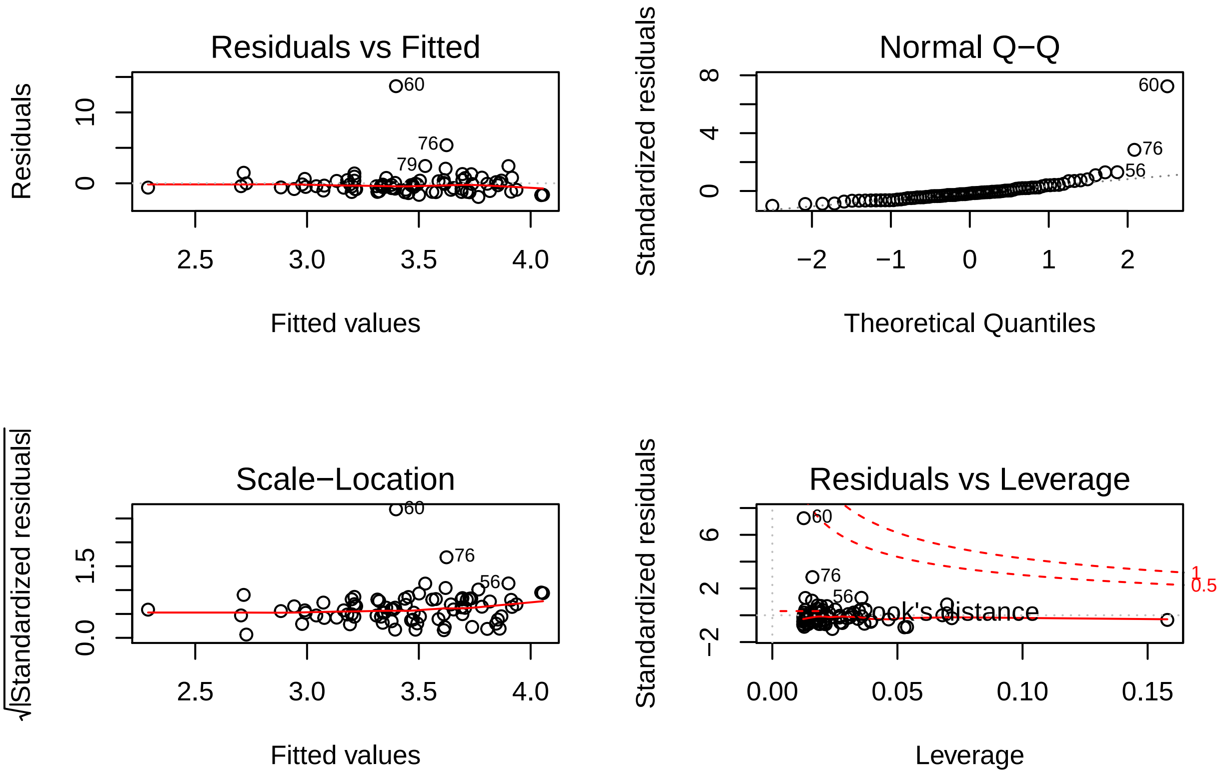
To begin, we analyzed whether this relatinoship is linear, or indeed strong, using a scatterplot and correlation.

```
## [1] -0.1755716
```



No evidence of a linear relationship appeared, based on a low correlation and a lack of a clear trend in the scatterplot. We nevertheless used a least squares regression hypothesis test to analyze this relationship. H_0 was that there is no relationship between stomach mean pH and GE time. H_A was that there is such a relationship. $\alpha = 0.05$ was our significance parameter.

```
##
## Call:
## lm(formula = GE.Time ~ S.Mean.pH, data = smart_pill)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9262 -0.8296 -0.3270  0.3683 13.7021
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.6376     0.7844   5.912 8.17e-08 ***
## S.Mean.pH    -0.3961     0.2499  -1.585    0.117
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.904 on 79 degrees of freedom
##   (14 observations deleted due to missingness)
## Multiple R-squared:  0.03083,    Adjusted R-squared:  0.01856
## F-statistic: 2.513 on 1 and 79 DF,  p-value: 0.1169
```

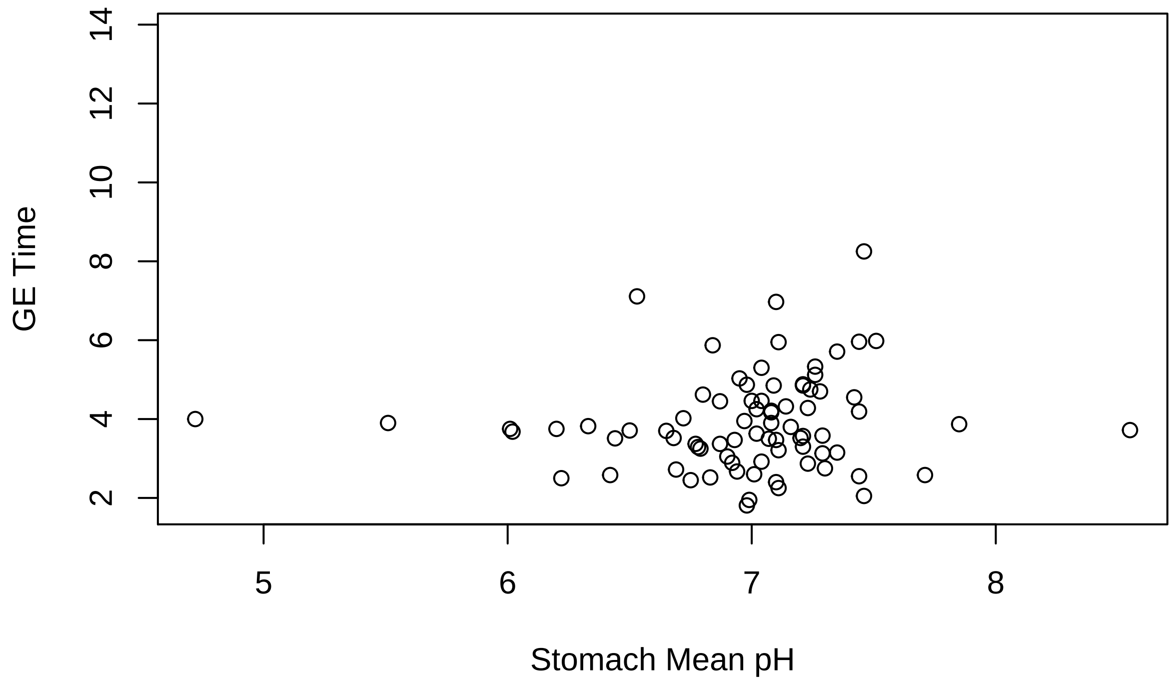


This test returned a p-value of $P = 0.117$, so $P > \alpha$ and we could not reject H_0. Furthermore, the residuals do not fit the assumed normal distribution thanks to a few extreme outliers (as evidence by their Q-Q plot above), which calls the value of this model into question. There is no statistically significant evidence for a relationship between stomach mean pH and GE time.

Small Bowel pH and Small Bowel Time

Below is the correlation and scatterplot of small bowel mean pH and SB time. These metrics did not show any evidence for a linear relationship, but we ran least squares regression analysis.

```
## [1] 0.1071285
```



H_0 was that there is no relationship between small bowel mean pH and small bowel transit time. H_A was that there is such a relationship. $\alpha = 0.05$ was our significance parameter.

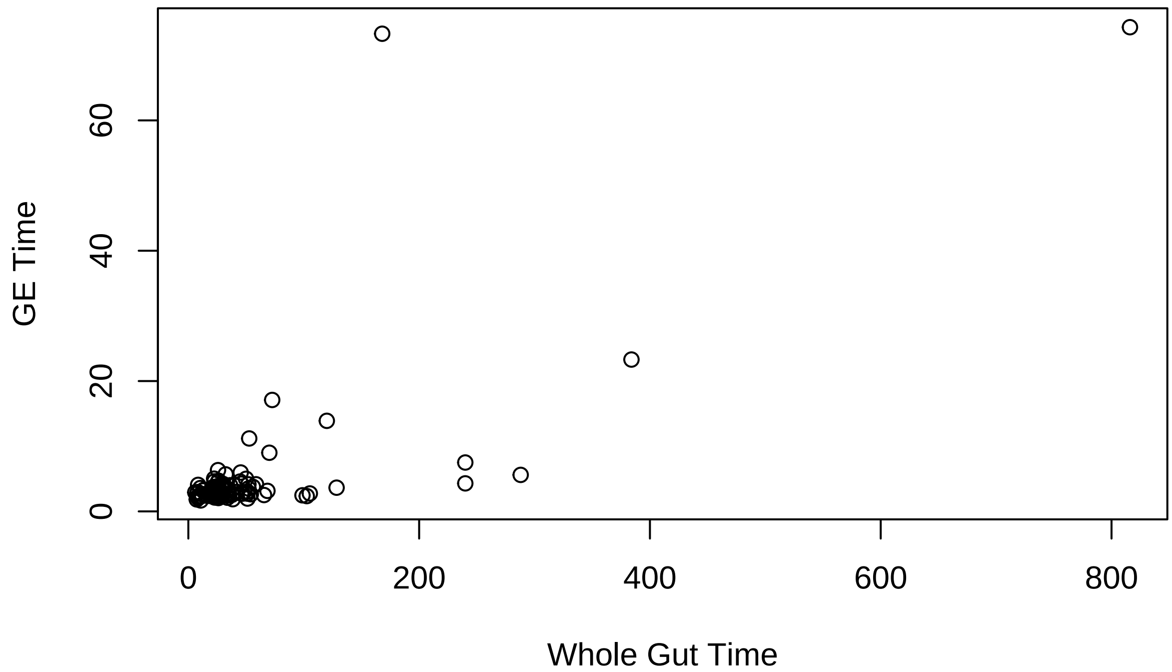
```
##
## Call:
## lm(formula = SB.Time ~ SB.Mean.pH, data = smart_pill)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.1010 -0.8490 -0.1244  0.6250  4.2131
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.0794     1.9420   1.071    0.288
## SB.Mean.pH    0.2624     0.2775   0.945    0.347
##
## Residual standard error: 1.217 on 77 degrees of freedom
##   (16 observations deleted due to missingness)
## Multiple R-squared:  0.01148,    Adjusted R-squared:  -0.001361
## F-statistic: 0.894 on 1 and 77 DF,  p-value: 0.3474
```

This test returned a p-value of $P = 0.347$. Because $P > \alpha$, we could not reject H_0, so there was no statistically significant evidence for a relationship between SB mean pH and SB transit time.

Whole Gut Time and GE Time

Below is the correlation and scatterplot for GE time and whole gut time.

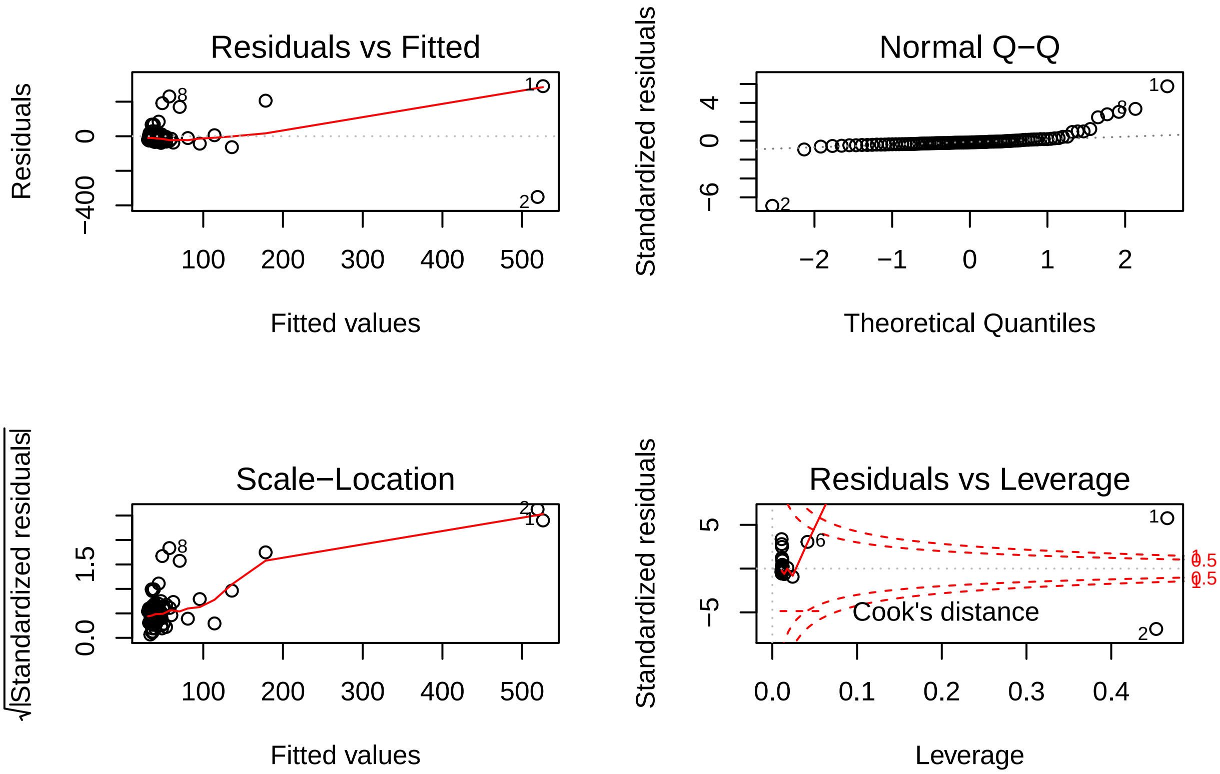
```
## [1] 0.7310908
```



The correlation between the two variables is reasonably strong, at $R = 0.7311$. However, it seemed from the plot that most of this correlation is due to a few outliers, more than a consistent linear trend. This relationship was still worth investigating.

We used a least squares regression hypothesis test to analyze the relatinoship between whole gut (WG) time and GE time. As before, H_0 was that there is no relationship between whole gut time and GE time. H_A was that there is such a relationship. $\alpha = 0.05$ was our significance parameter.

```
##
## Call:
## lm(formula = WG.Time ~ GE.Time, data = smart_pill)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -351.37  -22.44  -12.35    3.48  289.80
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  19.2612     8.1088   2.375   0.0197 *
## GE.Time       6.8228     0.6749  10.109   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 68.85 on 89 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.5345, Adjusted R-squared:  0.5293
## F-statistic: 102.2 on 1 and 89 DF,  p-value: < 2.2e-16
```



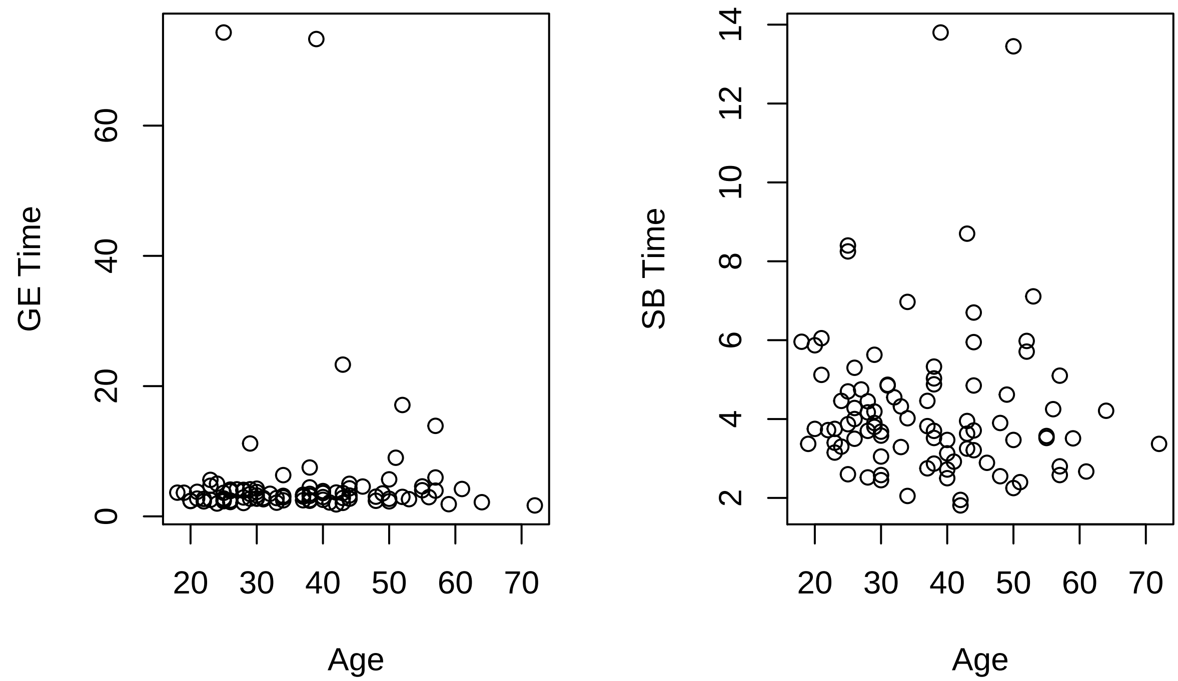
This hypothesis test returned a p-value indicating significance, since $P < \alpha$. However, this analysis was problematic.The residuals were not normally distributed, as evidenced by the Q-Q plot above. The founding assumptions of our hypothesis test were violated, so we could not reject H_0 or claim statistical significance in these results.

Age and Digestion Times

Below are the correlations for GE time and SB time, in that order. Below that are the scatterplots for Age and GE time (left) and age and SB time (right). In general, these plots and correlations made our hypothesized linear relationship seem extremely unlikely, but still ran our hypothesis test.

```
## [1] -0.01258527
```

```
## [1] -0.04050838
```



We used a multiple regression hypothesis test to analyze the relatinoship between age and our two response variables, GE time and SB time. As before, H_0 was that there is no relationship between age and the response variable. H_A was that there is such a relationship. $\alpha = 0.05$ was our significance parameter.

```
##
## Call:
## lm(formula = Age ~ GE.Time + SB.Time, data = smart_pill)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -18.761  -9.623  -0.250   7.166  34.550
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 38.355280   3.249187  11.805   <2e-16 ***
## GE.Time      0.009433   0.144098   0.065    0.948
## SB.Time     -0.273248   0.782892  -0.349    0.728
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.12 on 87 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.00169,    Adjusted R-squared:  -0.02126
## F-statistic: 0.07364 on 2 and 87 DF,  p-value: 0.9291
```

Both of the p-values were $P > 0.7$. We could not reject H_0. There is no relationship here, as expected.

Reproducing Original Results

In general, our attempts to find novel relationships in the paper's data were fruitless. We then attempt to reproduce the results of the original paper (that there is significant difference between the GE Time and SB Time of healthy patients and those of ICU patients with intracranial hemorrhage) using our own statistical methods.

We examined our response variables (GE time and SB transit time) compared to the explanatory variable of Group (healthy volunteers are in one group, critically ill patients with intracranial hemorrhage are in the other), which is categorical, by using the Student t-test. These tests are below.

Health Status and GE Time

Our first reproductive hypothesis test examined whether there is a statistically significant relationship between health status and GE time. Specifically, we performed a two-sample t-test.

For this test let μ_i be the mean gastric emptying time for ICU patients, μ_h be the mean gastric emptying time for the healthy volunteers, and $\Delta\mu$ be the difference between the two means.

H_0: $\Delta\mu = 0$.

H_A: $\Delta\mu > 0$.

We used $\alpha = 0.05$ as our significance parameter.

Before we could perform the hypothesis test, we determined whether to use the pooled or unpooled standard error estimate.

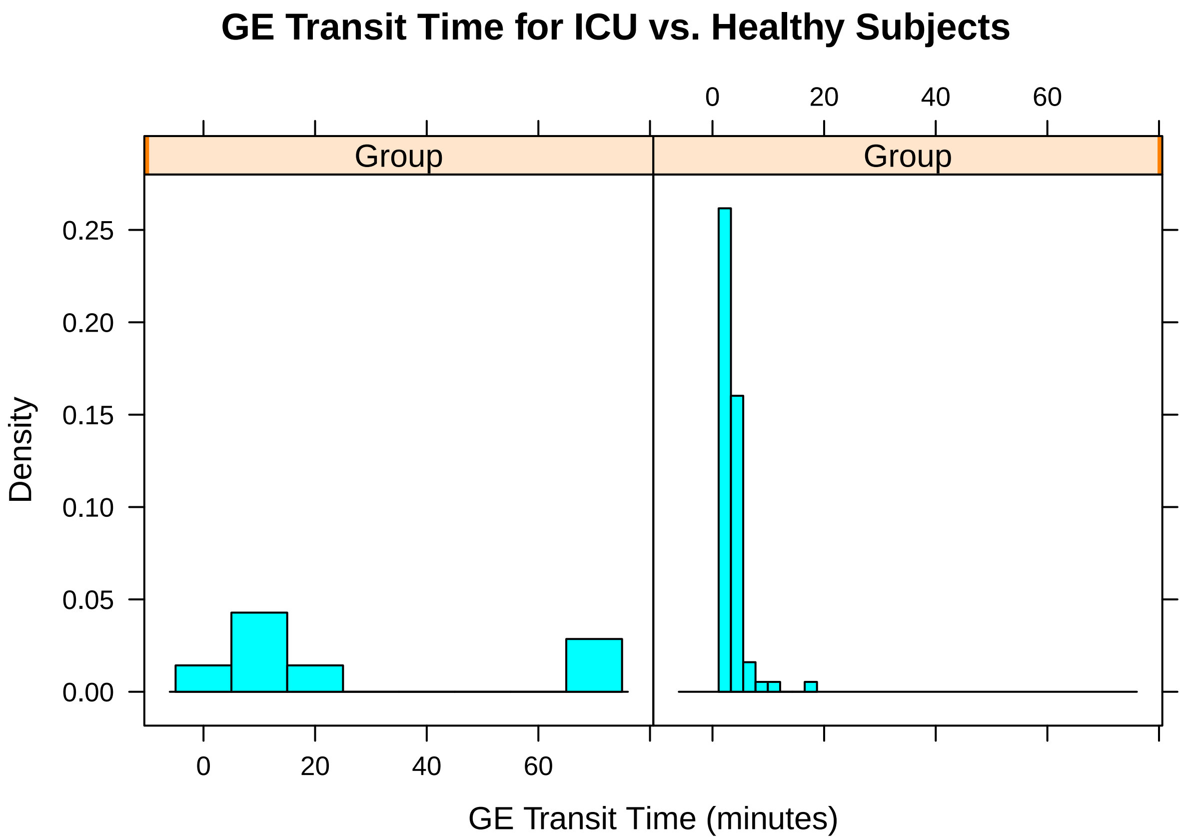
```
##   Group  min   Q1 median    Q3  max      mean        sd  n missing
## 1     0 4.30 6.55  13.90 48.30 74.3 28.885714 31.344560  7       1
## 2     1 1.68 2.51   2.99  3.88 17.1  3.532471  2.058441 85       2
```

Because the ratio of the standard deviations of the two groups ($\frac{31.34}{2.06}$) is greater than 2, we used the unpooled sample standard error estimate.

```
##
##  Welch Two Sample t-test
##
## data:  GE.Time by Group
## t = 2.1397, df = 6.0043, p-value = 0.03808
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  2.330974      Inf
## sample estimates:
## mean in group 0 mean in group 1
##       28.885714        3.532471
```

This test returned a test statistic of $t = 2.14$ and a p-value of $P = 0.038$. $P < \alpha$, so we rejected the null hypothesis. There is therefore statistically significant evidence that critically ill patients with intracranial hemorrhage have a longer gastric emptying time than healthy volunteers.

The histograms below show the distribution of GE time for critically ill patients (left) and for healthy patients (right). With such a small sample size for the ICU patients ($n = 7$), a few outliers in the ICU data may have altered the result of the hypothesis test. We therefore advise caution to readers of this study as they attempt to draw conclusions, and implore future researchers to repeat this study with a larger ICU patient sample size.



Health Status and SB Time

Our second hypothesis test examined whether there is a statistically significant relationship between health status and SB transit time. Specifically, we performed a two-sample t-test.

For this test let μ_i be the mean small bowel transit time for ICU patients, μ_h be the mean small bowel transit for the healthy volunteers, and $\Delta\mu$ be the difference between the two means.

H_0: $\Delta\mu = 0$.

H_A: $\Delta\mu > 0$.

We used $\alpha = 0.05$ as our significance parameter.

Before we performed the hypothesis test, we determined whether to use the pooled or unpooled standard error estimate.

```
##   Group  min   Q1 median   Q3   max     mean       sd  n missing
## 1     0 3.40 4.40   6.70 8.55 13.80 7.114286 3.618208  7       1
## 2     1 1.81 3.14   3.75 4.66 13.45 4.058916 1.612789 83       4
```

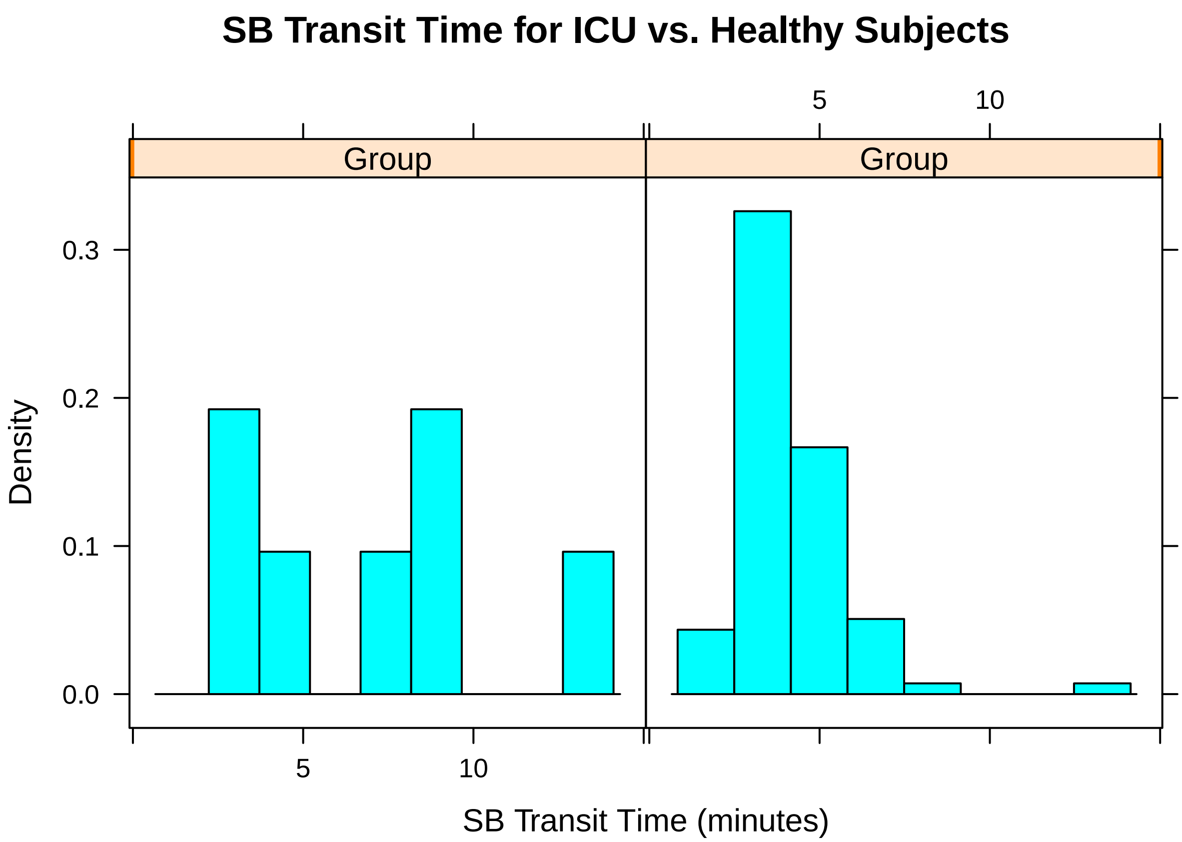
The ratio of the sample standard deviations ($\frac{3.62}{1.61}$) is about 2.25, so we again used the unpooled standard error estimate.

```
##
##  Welch Two Sample t-test
##
## data:  SB.Time by Group
## t = 2.2157, df = 6.2026, p-value = 0.03357
## alternative hypothesis: true difference in means is greater than 0
```

```
## 95 percent confidence interval:
##  0.3913902       Inf
## sample estimates:
## mean in group 0 mean in group 1
##        7.114286        4.058916
```

The hypothesis test returned a test statistic of $t = 2.22$ and a p-value of $P = 0.034$. Because $P < \alpha$, we can reject the null hypothesis. There is statistically significant evidence that SB transit time is longer in critically ill patients with intracranial hemorrhage than it is among healthy volunteers.

The histograms below show the distribution of SB transit time for critically ill patients (left) and healthy volunteers. We must advise similar caution to anyone wishing to draw strong conclusions from this hypothesis test. Again, the critically ill patient sample size is so small that a few outliers can skew the data. However, as you can see from the plots, the critically ill patient SB transit time does not have the extreme outliers present in the GE time for the critically ill patients. We again encourage other researchers to investigate this effect with a larger sample of critically ill hemorrhage patients.



Concluding Remarks

Our novel hypothesis tests returned little of interest. The only statistically significant relationship that our hypotheses turned up was one between GE time and whole gut time, but that linear regression model violated a few of its founding assumptions, so it didn't allow us to conclude anything meaningful. Further, that relationship is almost tautological, so in addition to being statistically dubious, it is practically uninteresting.

The paper set out to examine two questions: whether the wireless motility capsule is an effective way to collect gastrointestinal data for intubated patients, and whether critically ill patients had a significant decrease in metabolic activity. The answer to the first question was an emphatic yes. The answer to the second question was a yes up to statistical significance for a small sample.

We were able to reproduce the original paper's results with significance parameter $\alpha = 0.05$. We found p-values around $P = 0.03$. However, due to the extremely small sample size of the critically ill patient group ($n = 7$ in each case), we recommend caution in drawing strong conclusions from the data and recommend further studies with larger sample sizes.

Citations

Rauch et al. "Use of Wireless Utility Capsule to Determine Gastric Emptying and Small Intestinal Transit Times in Critically Ill Trauma Patients". Journal of Critical Care 2012; 27(5): 534.e7 – 534.e12.

https://www.ruf.rice.edu/~bioslabs/tools/stats/ttest.html

http://mathworld.wolfram.com/FishersExactTest.html

https://www.stat.auckland.ac.nz/~wild/ChanceEnc/Ch10.wilcoxon.pdf