**Project 1**

**Submission Rules**

1. This is an **individual project**. While you are welcome to ask for help from the instructor(s) and teaching assistants, you are expected to complete the data analysis and write-up of the report on your own.
2. We recommend you submit your document as a single **Word or PDF file.**
3. In your report, share statistical summaries or inferential results (**edited into tables**) or graphs that further your argument, giving titles to all tables and graphs/figures (example: Table 1 or Figure 1) for easy reference. Make sure you label graph axes and round statistical summaries or inferential results appropriately.
4. Write your report as you would for a client or collaborator, in full sentences and paragraphs. Make sure your presentation of your work is clean, readable, and professional. Sloppy presentation makes any data analysis less trustworthy. Points will be deducted for excessively sloppy submissions.
5. Your report should be **no longer than 2 pages** (not including appendix, tables and figures).
6. Your software code must be submitted as a separate file (\*.R, \*.Rmd, or \*.SAS).
7. Consult the grading rubric in the assignment in Canvas for more details on how the project will be graded.

**Background**

It is well established that inadequate sleep, whether from prolonged duty or circadian rhythm disturbances, degrades performance. Because there is no reason to assume that hospital personnel are immune to the performance degrading effects of sleep deprivation, resident work hours are increasingly being restricted to reduce fatigue and the potential for related errors. Even excluding the obvious sleep deprivation associated with overnight work, hospital personnel are likely to become progressively fatigued and work less effectively during the course of a normal workday. Anesthesiologists may be at particular risk because prolonged monitoring is especially impaired by fatigue. It is similarly likely that hospital personnel become progressively fatigued as the normal workweek progresses from Monday to Friday. An additional time-related factor that might influence clinician performance is that most new residents enter teaching hospitals in July and August, and the responsibilities of existing residents often precipitously increase at the same time. Long learning curves associated with anesthesia and surgical procedures may increase risks in the operating rooms during these months and therefore worsen patient outcomes.

This study (ref. 1) therefore tested the hypotheses that the risk of 30-day mortality associated with elective general surgery: 1) increases from morning to evening throughout the routine workday; 2) increases from Monday to Friday through the workweek; and 3) is more frequent in July and August than during other months of the year. As a presumed negative control, the investigators also evaluated mortality as a function of the phase of the moon. Secondarily, they evaluated these hypotheses as they pertain to a composite in-hospital morbidity (complication) endpoint. We will focus on this secondary endpoint for Project 1.

The data and data documentation for this study are found on Canvas, titled *Surgery Timing.csv* and *Surgery Timing Data Dictionary.docx*, respectively.

*Reference 1*: Sessler et al. “Operation Timing and 30-Day Mortality After Elective General Surgery”. *Anesth Analg* 2011; 113: 1423-8.

**Your Task**

The research questions are:

1. What is the effect of time of day on surgical in-hospital complications?
2. What is the effect of day of week on surgical in-hospital complications?
3. What is the effect of month on surgical in-hospital complications?
4. What is the effect of phase of the moon on surgical in-hospital complications?

*Considerations*

As the analyst, you will needto make many choices about how to approach this task. Things that you can consider:

* How many models are needed to answer the question
  + Should there be a single model with all of the predictors of interest or should there be different models for each predictor?
* How the predictors are coded to meaningfully answer the question
  + Should the variables be left as they were collected or should some of the categories in a predictor be combined to ease interpretation?
* Which variables (if any) are adjusted for in the model(s)

Whatever you decide, be sure to state clearly about why you made particular choices

**Temporal Factors of In-Hospital Surgical Complications**

2/22/2024, Kath Fillman

**Introduction.** This is an analysis of the effect of an elective general surgery’s timing in relation to hour, day of the week, month, and moonphase on if an in-hospital surgical complication will arise.

**Dataset.** The dataset ‘Surgery Timing.csv’ was provided by the instructor and is a cleaned version of the data from Sessler et al. “Operation Timing and 30-Day Mortality After Elective General Surgery”. *Anesth Analg* 2011; 113: 1423-8. It contained information from 32001 patients’ elective general surgeries which included 25 variables which includes a binary variable indicating if there was an in-hospital surgical complication and information about the surgery’s timing (hour, day of week, month, and moon phase). The variable hour was adjusted to include only the greatest integer as opposed to both hour and minutes; this was done to ease interpretation.

**Exploratory Data Analysis.** Of the 32001 patients included in this analysis, 4264 of them experienced in-hospital surgical complications. Figure 1A shows the distribution of the hour of operation grouped by the presence of a complication, the most common time of operation for both groups was within the hour of 7am. The proportion of in-hospital complications fluctuated considerably throughout the day and were higher within the hours of 7am, 9am, 10am, 2pm, 3pm and 4pm. The variable containing the day of the week the surgery was performed on (dow) includes only weekdays (Monday-Friday). Figure 1B shows the distribution of the day of the week the surgeries took place grouped by if there was an in-hospital complication. The proportion of in-hospital complications is consistent throughout the week despite the number of surgeries fluctuating (table 1). Figure 1C shows the distribution of the month of operation grouped by if there was a complication present. The proportion of in-hospital complications is consistent throughout the months despite the number of surgeries fluctuating (table 1). Figure 1D shows the distribution of the moonphases grouped by if there was an in-hospital surgical complication. The proportion of in-hospital complications is consistent throughout the moon phases despite the number of surgeries fluctuating (table 1).

**Methods.** An association between in-hospital complications and the surgery’s time of day, day of the week, month, and moon phase was quantified using logistic regression. A logistic model was chosen as the outcome of interest was binary (presence of a surgical complication). Two models were constructed using the glm() function from the R jtools package. One with all variables and one with only the statistically significant predictors; these will be compared later to choose the better fitting model. Creating one model was chosen over multiple, individual models for each predictor as it delivers the greatest accuracy in its prediction of in-hospital complications while adjusting for temporal factors. P-values less than 0.05 were considered significant. All analyses were performed in R-studio version 2023.12.1, build 402 and R version R-4.1.3.

**Results.** The results of the logistic regression model are shown in table 2. The predictors hour and day of week were both found to be statistically significant with p values of 0.0002 and 0.0045 respectively. Predictors month and moon phase were not found to be statistically significant with p values of 0.6226 and 0.5085 respectively. The first model created using table 2 is log(odds) = 0.1163 + 1.0216(hour) + 1.334(day of week) + 0.9976(month) + 0.9903(moon phase). The second model created is log(odds) = 0.1140 + 1.0207(hour) + 1.333(day of week). A likelihood ratio test was performed and found that the models fit equally well (p-value of 0.7167), thus the less complex model was chosen. Likelihood ratio tests between models with singular predictors and the second model were also performed (p-value of 0.0047 for hour and 0.0002 for day of week) and found that the model with two predictors was a better fit. A Hosmer and Lemeshow goodness of fit (GOF) test returned a p-value of 0.0.0298, indicating that this model is not a good fit for the data. This model has an AIC of 25106.1488, which is very high and corroborates poor model fit.

**Conclusion.** After testing for model fit, it does not seem that temporal factors are meaningful predictors of in-hospital surgical complications. Given their statistical significance in the model, variables such as the hour of operation and say of the week may be predictors of in-hospital complications, however due to poor model fit it is unlikely they are meaningful. Based on the model, the odds of surgical complication are increased by 2.1% every hour after 6am when surgeries begin each day). Likewise, the odds ratio for the day of the week is 2.81 times the day prior. The temporal factors hour of operation, day of the week may have an effect on in-hospital surgical complications while variables month and moonphase do not.

**Appendix.**

**Table 1. Number of In-Hospital Complications Based on Day of the Week, Month, and Moon Phase**

|  | | **Number of In-Hospital Complications** | |  | | **Number of In-Hospital Complications** | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **No** | **Yes** | **No** | **Yes** |
| **Hour** | 6 | 514 | 48 | **Month** | January | 2263 | 407 |
| 7 | 9290 | 1341 | Febuary | 2163 | 343 |
| 8 | 3236 | 571 | March | 2373 | 324 |
| 9 | 1446 | 218 | April | 2377 | 321 |
| 10 | 2204 | 297 | May | 2321 | 333 |
| 11 | 2492 | 363 | June | 2584 | 410 |
| 12 | 2359 | 404 | July | 2024 | 301 |
| 13 | 2294 | 329 | August | 2715 | 462 |
| 14 | 1775 | 288 | September | 2784 | 424 |
| 15 | 1066 | 201 | October | 2312 | 377 |
| 16 | 631 | 114 | November | 2219 | 325 |
| 17 | 295 | 61 | December | 1602 | 237 |
| 18 | 134 | 29 | **Moon Phase** | New Moon | 6641 | 1067 |
| 19 | 1 | 0 | First Quarter | 7067 | 1033 |
| **Day of the Week** | Monday | 6097 | 908 | Full Moon | 6963 | 1088 |
| Tuesday | 9796 | 812 | Last Quarter | 7066 | 1076 |
| Wednesday | 5339 | 927 | **Total** | | 27737 | 4264 |
| Thursday | 4841 | 794 |  | | | |
| Friday | 5264 | 823 |

**Table 2. Model Summaries**

| **Model 1** | | | |
| --- | --- | --- | --- |
|  | **Log(odds)** | **95% Confidence Interval** | **p-value** |
| **Complication (intercept)** | 0.1163 | 0.0985, 0.1373 | N/A |
| **Hour** | 1.0208 | 1.0098, 1.0318 | 0.0002 |
| **Day of the Week** | 1.0334 | 1.0102, 1.0571 | 0.0045 |
| **Month** | 0.9976 | 0.9880, 1.0073 | 0.6226 |
| **Moon Phase** | 0.9903 | 0.9620, 1.0194 | 0.5085 |
| **Model 2** | | | |
|  | **Log(odds)** | **95% Confidence Interval** | **p-value** |
| **Complication (intercept)** | 0.1140 | 0.0999, 0.1301 | N/A |
| **Hour** | 1.0207 | 1.0097, 1.0317 | 0.0002 |
| **Day of the Week** | 1.0333 | 1.0101, 1.0570 | 0.0047 |

**Figure 1. Distributions of Variables of Interest**

