JOURNAL OF
EMERGING INVESTIGATORS

Article

Predicting the factors involved in orthopedic patient hospital stay

Rahul D'Souza[1], Shaan Kamal[2]
[1]Sanford High School, Hockessin, Delaware
[2]University of Connecticut, Storrs, Connecticut

SUMMARY

Long hospital stays can be stressful for the patient for many reasons. Patient length of stay also concerns the hospital from a business standpoint who would like to minimize the length of a patient's stay, without compromising care. In this study, we investigated what factors are associated with length of hospital stay among orthopedic surgical cases. We hypothesized that age would be the greatest predictor of hospital stay among patients who underwent orthopedic surgery. We also hypothesized that the length of stay would, on average, range from one to three days. We used a dataset from the New York Statewide Planning and Cooperative System which comprises a group of hospitals in New York City. Using machine learning models to predict a patient's length of stay, we employed exploratory data analysis, support vector machines, principal component analysis and random forest models. Through our models, we found that severity of illness was indeed the highest factor that contributed to determining patient length of stay. The other two factors that followed were the facility that the patient was staying in and the type of procedure that they underwent.

INTRODUCTION

Predicting the length of hospital stay for orthopedic patients is a valuable indicator for many stakeholders, such as the patient, the patient's family, hospital staff, and hospital management. For all involved, a shorter length of stay is preferred. Length of stay has been associated with increased complications, readmissions, and decreased patient satisfaction. Hospitalization demonstrably exacerbates patients' emotions and increases feelings of depression and anxiety (1). The average cost for a three day stays at the hospital is approximately $30,000 (2). Extended stays can incur many additional costs for the patient. Medical debt itself can poorly affect health outcomes (3). In addition, if hospital stays were decreased, hospital staff would be able to treat additional patients as well as hospital management that have finite resources, such as beds, equipment, and doctors, would get to utilize them more efficiently.

In addition, better and faster treatment would increase patient turnaround and therefore increase hospital revenue (4). Hospital management is also looking for ways to increase the stature and recognition of their hospital, as it will attract more patients. All of this can be greatly improved by shorter hospital stays. For the patient and their family, a shorter length of stay leads to a better disposition, less exposure to pathogens like COVID, less costly and better expectation of how long they will be out of work or away from family. The outcome of a shorter stay is a healthier and happier patient. Shorter length of stay may also predict a more positive outcome with the patient's recovery time or outcome results, as longer lengths of stay increase the patient's exposure to hospital germs (5). The bottom line is that a shorter length of stay is much better than a longer length of stay, for all parties involved.

With this research, we have been able to identify who stays longer. This is an important first step into further dissecting why patients stay longer and to find effective solutions to shorten their hospital stay. Given that facility ID is a highly contributing factor, this is something that can be explored further to break down the key aspects of each facility to identify possible reasons for increased length of stay. Perhaps pulling out facility zip code, number of doctors per patient, number of nurses per patient, age of equipment as well as other potential facility attributes, would identify strong indicators. Further research in this area can go into the differences between each facility and point out shortcomings in decreasing the length of time a patient stays in the hospital.

RESULTS

Using the SPARCS dataset, we looked at patient data such as, location of hospital, severity of illness, age, gender, and race. Using machine learning methods, such as support vector machines and principal component analysis, we were able to cluster the large amount of data into groups and eliminate any data that we did not find useful. We determined younger patients typically were less likely to undergo orthopedic surgery, with 86.48% of all surgery patients being above 50 years old. There were over 800 patients who were in the age range of 50 years and older, whereas there were

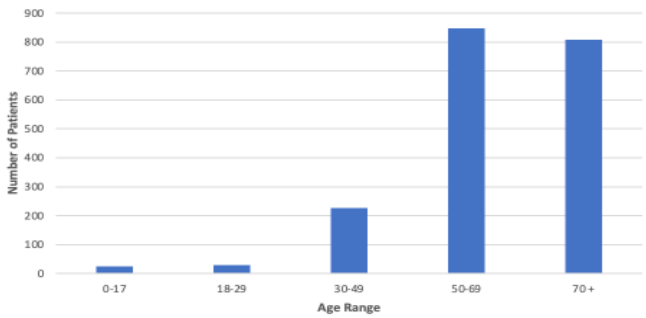

Figure 1: Patient ages included in the SPARCS data set that were pulled out for Orthopedic cases. Number of patients from the data set in New York hospitals stratified by age group. The majority of patients from the study were in the age range of 50-69 years old.

DOI: https://doi.org/10.59720/22-036

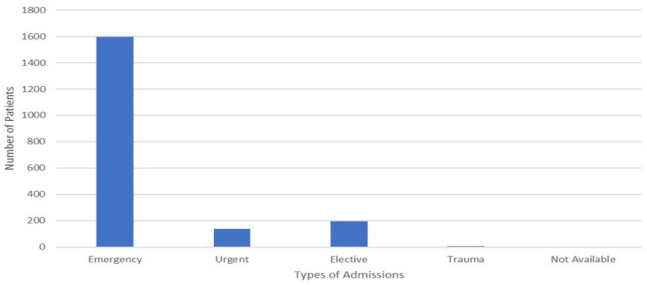

Figure 2: Type of orthopedic admissions in New York Hospitals. Most patients from the dataset were emergency admissions, which meant that these admissions were not scheduled. When doing an operation that is scheduled, doctors consider the patient's recovery period. For procedures that were emergent, the doctors are not aware of the case they are treating.

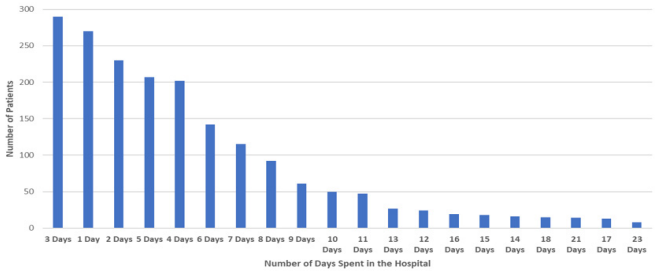

Figure 4: Number of days orthopedic patients spend in New York hospitals. Most of the patients from the dataset were admitted only for 3 days. The majority of hospital stays for orthopedic cases are typically under 10 days as indicated on the graph. The graph is skewed right due to the lack of patients staying in the hospital over 10 days.

less than 100 total patients in the 0–17 and 18–29 age range (**Figure 1**). Eighty-two percent of the patients in our dataset came in for emergency orthopedic surgery while the elective and urgent cases comprised 18% of the patients (**Figure 2**). The number of people who were admitted for emergencies was above 1400 whereas urgent and elective admissions were both below 200 (**Figure 3**).

We next analyzed the length of stay for all patients. We saw that three days was the most common hospital stay length in the New York hospitals (**Figure 4**). Most patients stayed for one to three days, and the maximum length of stay started to dramatically drop off at around 10 days. All the lengths of stays from 1–10 days were in the top 10 of our data count, whereas the larger outliers of 15 days or higher had much fewer counts. We calculated the relative importance of our factors that we researched, by comparing each case with all their unique factors from the dataset and seeing which ones correlated with each other. The highest relative importance was determined by which factor had the highest coefficient of determination or correlation. We calculated the coefficient of determination, or the percentage of data points that fall on the line on the graph, by taking the sum of the distance of the data from the mean, and it was calculated 0.15 (1). A coefficient of determination or R2 of 0.01 is a weak correlation; a coefficient of determination of 0.09 is a moderate association; and one that is 0.25 or bigger is a strong correlation (1). Severity of illness code had the highest coefficient of determination or correlation with patient length of stay (**Figure 5**). From our research, we found that severity of illness was the most important factor in determining how long a patient stayed in the hospital after their orthopedic surgery. The CCS procedure code which explains what surgery the patient had was another factor that predicted length of stay. Other factors like the type of facility, facility ID, zip code, disposition and age, also had a varying effect on the patient's length of stay.

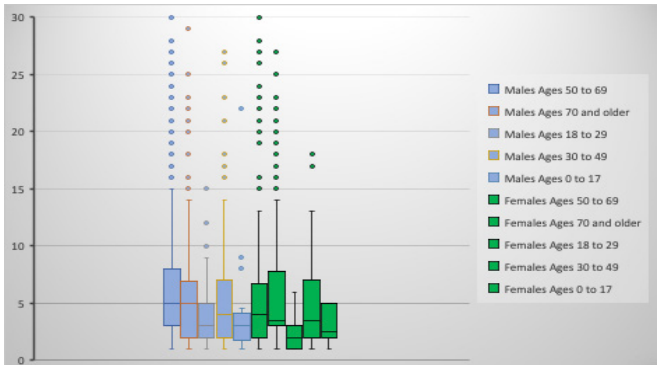

Figure 3: Correlation between age, gender and length of stay for orthopedic patients in New York hospitals using boxplots. The boxes represent the upper and lower quartiles, the lines passing through each box is the median age for each gender of patient, and the lines coming from the top and the bottom, the whiskers, take into account the minimum and maximum from each range.

DISCUSSION

Knowing the factors associated with length of hospital stay can be very helpful in managing the efficiency of hospitals, and ultimately lowering costs for hospitals and patients.

The length of stay is very important not only for a patient's health, but also financially to the hospital and doctors that are treating them. Our study highlights how important it is for further research to be done in this area. It can help hospital systems and physicians drive better outcomes, give them better ratings, and attract more patients. Insurance and other health care companies have been shifting more of the financial risk of patients onto the providers to improve the quality of health care (6). All of this in turn will reduce overall costs for the patients, hospitals, and doctors, while driving increased revenue for the hospitals and doctors. Those longer stay outlier patients would be costly for the insurance companies and the hospital systems. In addition, those patients would be

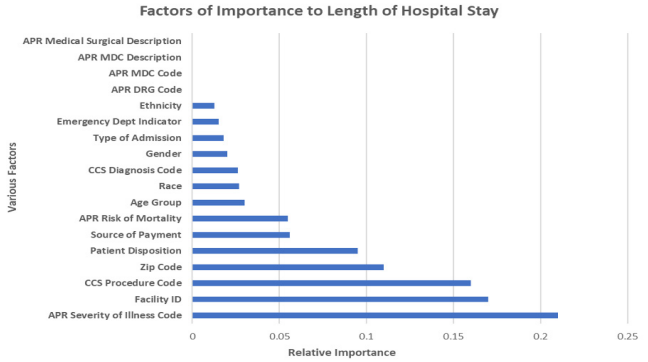

Figure 5. Factors that determine the length of stay. The severity of illness affected the patient's length of stay the most followed by the facility where the patient was staying at and the type of procedure that the patient underwent. Another factor that is important to note is that ethnicity played a role in how long the patient was admitted for.

DOI: https://doi.org/10.59720/22-036

detrimental for the reputation of the physicians and hospital systems involved. Further expanding on this research, we could possibly add in the factor of whether patients had health insurance or not to our data. If a patient does not have health insurance, they will have a shorter length of stay to avoid costs. This will not only make the patient have less to worry about regarding insurance, but also will be beneficial to the hospital in terms of increasing their capacity.

Another modifying factor is the CCS procedure code that explains what surgery the patient had. Other factors like the type of facility, facility ID, zip code, disposition and age, also had a varying effect on the patient's length of stay. Some other interesting factors that determined the length of stay were the facility ID and zip code (**Figure 5**). These factors show that individual hospitals in the NYC hospital system have better outcomes. This could affect possible future patient choice as to where they would want to be treated. This signifies that further research in this area is needed. In addition, government and insurance payment structures could use this data to incentivize or penalize hospitals that have better or worse outcomes. An additional key factor in determining the length of hospital stay is the CCS procedure code, as that tells us what surgery the patient had. This data can be used to determine what surgery or type of surgery will provide the best outcome and reduced length of stay for a patient. For example, a patient may have a hip fracture and a different procedure or surgical code will reduce length of stay even for the same injury. One of the more interesting aspects of our research is that many factors may either overlap or combine to increase the length of stay. A severe fracture in a young child might lead to a one- or two-day hospital stay, whereas a severe hip fracture in a 90-year patient might lead to a five to seven day stay. Finally, there are other factors, such as race, gender, and ethnicity, that play a role in the length of stay as we found by analyzing our dataset. These are factors that we feel should not play as much of a role in the length of stay as we found; however, there could be other environmental or societal challenges that may be playing a role. For example, a study from 2017 focused on which patient's specific characteristics affected the patient length of stay the most (7). They found that congestive heart failure, patients who were classified as underweight, patients who had diabetes, and patients who had morbid obesity all saw an increase in length of stay (7). Also, another study in 2021 was conducted at Seti Provincial Hospital in Nepal. Researchers conducted this study on 800 patients who were admitted there for an orthopedic related injury. From the data they collected, they concluded that their average length of hospital stay among the patients was 2.87 days (8). This also supported our research.

In conclusion, it is very possible to do further study on each facility, taking this research a step further in addition to looking at other factors that contributed to a longer length of stay for the patient.

MATERIALS AND METHODS

We used a dataset from the New York Statewide Planning and Cooperative System (SPARCS) which comprises a group of hospitals in New York City. SPARCS is an all-payer healthcare data reporting system used to collect patient-level detail for each hospital inpatient stay, hospital outpatient visit, and associated hospital clinic visits. The SPARCS dataset included all orthopedic surgery patients, their length of stays, age ranges, socioeconomic status, type of injury, type of surgery, and several other important factors that would affect their length of stay. The multitude of factors in this dataset made it a valuable choice for our research.

In this study, we sought to determine the important factors that contribute to length of stay and the degree to which they affect patients. We wrote code using Python to create various machine learning models which used patient information to extrapolate a patient's length of stay. Before writing the code, the dataset was imported and any extraneous factors were removed. Some of the factors that were removed were total charges, total costs, admit day of week, facility name, discharge year, and discharge day of week. We imported Pandas due to our research extracting and manipulating data to create tables and graphs. The machine learning models we employed included EDA (exploratory data analysis), support vector machines, and principal component analysis. The EDA tools helped us create the correlation to determine which factor was the most important. By using EDA, we were able to see trends and patterns in our data regarding patient length of stay. We used it to determine the correlating factors of patient length of stay and we were able to create graphs using this data (1). Support vector machines and principal component analysis were added to the code to help us cluster the large amount of data into groups and get rid of any clusters that were not pertinent to our study.

We worked to find out which factors had the most correlation with longer length of stay and whether our assumptions were correct. We concluded that severity of injury and facility ID were the most important factors in determining patient length of stay in addition to the procedure code.

ACKNOWLEDGEMENTS

We would like to thank Dr. Deepika Thacker and Dr. Sanjay Aneja for looking over our project. From a medical professional's point of view, their support and comments were helpful during the editing process.

Received: January 17, 2022
Accepted: November 28, 2022
Published: December 13, 2023

REFERENCES

1. Alzahrani, Naif. "The effect of hospitalization on patients' emotional and psychological well-being among adult patients: An integrative review." *Applied Nursing Research*, vol. 61 (2021): 151488. https://doi.org/10.1016/j.apnr.2021.151488.
2. "Why Health Insurance is Important" *Healthcare.Gov*. www.healthcare.gov/why-coverage-is-important/protection-from-high-medical-costs/.
3. "How Medical Debt Affects Health" *Sycamore Institute TN*. www.sycamoreinstitutetn.org/wp-content/uploads/2021/05/2021.05.19-FINAL-How-Medical-Debt-Affects-Health.pdf. Accessed 2021 May
4. "Patient-Centered LOS Reduction Initiative Improves Outcomes, Lowers Costs" *Health Catalyst*. www.healthcatalyst.com/wp-content/uploads/2021/05/Patient-Centered-LOS-Reduction-Initiative-Improves-Outcomes-Saves-Costs.pdf. Accessed 2017
5. "Get a Faster Hospital in Five Days" *QI Macros*. www.

DOI: https://doi.org/10.59720/22-036

qimacros.com/pdf/A-Faster-Hospital-in-Five-Days.pdf. Accessed 2021 July
6. Williams J, et al. "The Association Between Hospital Length of Stay and 90-Day Readmission Risk Within a Total Joint Arthroplasty Bundled Payment Initiative." *Journal of Arthroplasty,* 2017;32:714-8. https://doi.org/10.1016/j.arth.2016.09.005.
7. Gholson, J, et al. "Patient Factors Systematically Influence Hospital Length of Stay in Common Orthopedic Procedures." *The Iowa Orthopedic Journal*, vol. 37 (2017): 233-237.
8. Joshi, Pramod. "Hospital Stay of Orthopedic Cases in a Tertiary Care Hospital: A Descriptive Cross-sectional Study." *JNMA*, vol. 59,236 310-312. 30 Apr. 2021, https://doi.org/10.31729/jnma.6400.