Cover sheet

NATIONAL UNIVERSITY OF SINGAPORE

SP1541 EXPLORING SCIENCE COMMUNICATION THROUGH POPULAR SCIENCE

Assignment: Pre-course Writing Task

Name: Claudeon Reinard Susanto

Matriculation No.: A0239079R

Tutorial Group: S19

Major(s): Data Science and Analytics

Discipline of the selected research article: Statistics

*Delete as appropriate

Part I: Science News Article

Using statistical modelling and big data to minimise wastage of blood platelet units in hospitals

By: Claudeon Reinard Susanto, 17 Aug 2022

Blood platelets serve an important function in our body; they cluster together and form a clot which will prevent and stop bleeding. For the same reasons, hospitals and other tertiary care facilities require a steady supply of blood platelets to reduce loss of blood in patients who are actively bleeding because of injury or surgery, as well as prevent bleeding in patients with low blood platelet count.

However, the demand for platelets is highly volatile. There might be a surge in demand at times due to natural disasters or mass accidents among many other factors. Hence, hospitals and other tertiary care facilities would always have a surplus of platelets ready at hand as they act as a necessary precaution to avoid shortage in case of these unforeseen circumstances.

Similarly, on some days, demand for platelet units would be much lower than the current stock of platelets ready at hand. Unlike products like canned food and drinks, blood platelets are perishable and they only have three days in which they are available for use. Therefore, these excess platelets are undeniably wasted.

In the US, this wastage amounts to almost US$80 million due to the difficulty and costs of collecting platelets from donors. This wastage represents the amount of time and money that could have been better spent on improving the quality and safety of other areas in the healthcare industry!

This issue of wastage stems from one problem: healthcare providers find it difficult to predict the usage of blood platelets on a given day. Usage of blood platelets is highly unpredictable, and thus it is hard to minimise wastage by reducing platelets stock without ensuring that at the end of the day, there would be enough blood platelets for everyone.

Hence, it is crucial to forecast the usage of platelets days in advance such that wastage can be minimised as much as possible, while at the same time ensuring that there are enough platelets ready at hand even when there are emergencies.

Previous studies had yet to develop an effective method for predicting the usage of platelets in hospitals. These studies relied upon a small number of patient population or used charts as a tool which would be inaccurate. Nevertheless, one study found that platelets usage is higher on weekdays as compared to weekends.

However, recently, researchers at Stanford University have successfully devised a mathematical model using big data and statistical modelling to predict platelet usage days in advance.

The assumption used by the researchers is that large amounts of data would increase the accuracy of prediction of platelet usage 3 into the future. This prediction model would then be used to guide hospitals on how to recruit donors such that the number of platelet units collected would match the demand 3 days in the future. In this case, predicting 1 day in advance is not enough as platelets need to spend 2 days in testing after collection from donors.

Different from previous studies, the Stanford University researchers utilised very large amounts of data from hospitals which are often called 'big data'. The big data contains 29 months of patients' Electronic Medical Records (EMR) from an associated hospital, Stanford Health Centre (SHC), as well as the total amount of expired and wasted platelets from Stanford Blood Center (SBC) where blood platelets are collected from donors.

Examples of EMR data collected from SHC include patients' platelet count and other blood count data. However, the researchers did not consider individual patient's data in their analysis; they considered the bigger picture, taking account only the total number of patients with abnormal blood count.

The researchers also collected the number of patients from different units in SHC, such as Trauma, Neurosurgery, and Vascular Care. Besides the day-of-week status (weekend or weekday), it was found that demand for platelets on a specific day was strongly influenced by the average demand for the previous 7 days.

After combining all the factors together (blood count, unit in the hospital, day-of-the-week, average number of platelet units transfused in the last seven days), there were 43 variables in total which could influence the platelet unit usage. These variables are then fed into the statistical tool IpSolve which determines the mathematical relationship between the 43 variables and usage of platelets. Hence, if the values of these 43 variables are known, we can accurately predict the demand for platelets for the next 3 days.

Using this model to guide donor recruitment strategy, the scientists were able to significantly reduce the initial platelet wastage rate of 10.5% to 3.2% in SHC over a period of 29 months. Furthermore, using this model, the hospital did not experience any shortage of platelets over the same period. If this approach is adopted nationwide, this could translate to cost savings of over US$80 million.

(797 words)

References:

Guan, L., Tian, X., Gombar, S., Zemek, A. J., Krishnan, G., Scott, R., Narasimhan, B., Tibshirani, R. J., & Pham, T. D. (2017). Big data modeling to predict platelet usage and minimize

wastage in a tertiary care system. *Proceedings of the National Academy of Sciences*, *114*(43), 11368–11373. https://doi.org/10.1073/pnas.1714097114

Part I: Science News Article (Revised)

Predicting platelet usage using statistics and big data

Scientists at Stanford successfully predicted blood platelet usage in hospitals with remarkable accuracy using statistics, potentially saving $80 million in health care costs.

By: Claudeon Reinard Susanto, 17 Sep 2022

Have you ever wondered how a small cut on your finger can heal in just a few days? It turns out that blood platelets present in our blood are responsible for healing injuries!

In the event of an injury, platelets will cluster together at the site of the wound. A clot will be formed, and this will prevent blood from leaking out. This clot also acts as a scaffolding where tissues and skin cells regrow, which is exactly the reason why wounds can heal very quickly.

Because of its ability, hospitals use platelets from donors to reduce blood loss in patients who are recovering from surgery and trauma, as well as prevent bleeding in patients with low platelet counts in their blood.

The problem is it is very hard to predict the daily demand for platelets in hospitals. Because platelets donated from blood donation centers need to spend two days in testing before being certified as usable, hospitals have to predict their platelet usage three days in advance.

Of course, any rational hospital would just stockpile platelets as much as they can to prevent a shortage in case there is an unforeseen surge in demand, which is usually caused by natural disasters or accidents. However, unlike canned food, platelets only have a very short shelf life of three days before expiring. These expired platelets end up being wasted.

Approximately 10.5% of 2 million platelet units are wasted annually, and this represents $80 million in healthcare costs! This could definitely be better used to improve the quality and effectiveness of other areas in the healthcare industry. An accurate method for prediction of platelet usage is consequently necessary to minimize wastage of platelets, while at the same time maintaining a buffer stock of platelet units as a safeguard measure.

Numerous studies have tried to predict platelet usage using patients' data. However, these studies had very limited sample sizes and so cannot be used as a general model for the larger population.

At Stanford University, researchers Leying Guan and colleagues from the Department of Statistics sought to solve the limitations in current forecasting methods. Their goal was to develop a mathematical model which would be able to forecast platelet usage three days in advance accurately.

Unlike previous studies, this time the Stanford researchers utilized large amounts of data collected over a very long period of time. They hypothesized that certain measurements in patients' data, such as platelet count and blood composition, greatly influence the usage of

platelet units. To test this claim, the researchers collected data from the university's associated hospital, Stanford Health Care (SHC), as well as the number of platelets transfused and wasted each day for more than 800 days.

Imagine you have to read textbooks for more than 800 days non-stop. You have to take a quiz at the end of each day, and the suggested answers as well as your quiz score will be given. You will surely learn from your mistakes and perform well for your final exam at the end of the period.

In the same way, our model learns platelet usage trends from the SHC data continuously for 800 days. The model is also tested every day to see if the prediction matches the actual amounts of platelets used. Over time, the gap between prediction and actual value will narrow. The result in the end was a model that can accurately predict platelet usage with little error. The prediction results -- the optimal quantity of platelet units that minimizes wastage -- would then be used to order platelets from blood donation centers.

At the beginning of the study, the hospital wasted 1 in 10 platelet units on average daily. Using this new model, only 1 in 30 platelet units were wasted and there were no days where there was a shortage of platelets. If this model is used nationwide, it could translate to cost savings of over 80 million dollars.

Despite the potential cost savings, this model is still limited in use as it is based on data from one hospital. Every hospital has different patient needs, so in order for the model to work in a different hospital, it needs to be trained again. But training the model means that data must be collected over long periods of time so that the model is as accurate as possible. It is also costly to collect and manage big data.

In spite of the limitations, this model still has the potential to save millions of dollars of costs if applied widely. Hopefully, in the future, you could see your neighborhood hospital using this mathematical model.

(799 words)

References:

Guan, L., Tian, X., Gombar, S., Zemek, A. J., Krishnan, G., Scott, R., Narasimhan, B., Tibshirani, R. J., & Pham, T. D. (2017). Big data modeling to predict platelet usage and minimize

wastage in a tertiary care system. *Proceedings of the National Academy of Sciences*, *114*(43), 11368–11373. https://doi.org/10.1073/pnas.1714097114

Hematology glossary. Hematology Glossary - Hematology.org. (n.d.). Retrieved September 11, 2022, from https://www.hematology.org/education/patients/blood-basics

Part II: Post-review Reflection

Reflection

As a peer reviewer, giving feedback to my peers has enabled me to identify certain strategies that work well, and to be aware of certain parts/moves that are missing or unclear. These have helped me identify missing and ineffective parts in my news article.

Regarding feedback given by my peers, they helped me see things in my article that I did not see such as the missing link in the introduction part of my article. However, I have difficulty adding more links for clarity as I have almost reached the word count limit. Nevertheless, I'm trying to improve on clarity and structure as I'm planning my new science news article.

All in all, the review experience was inspiring as I gained some insights by reading my peers' articles as well as the reviewers' comments. I will incorporate their strategies and suggestions into my new article.

(145 words)