**Revised Pre-course News Article**

**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.

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

This component is none other than blood platelets. In the event of an injury, platelets will cluster together at the site of the wound. When enough of them are present in one location, a clot will be formed, and this will prevent blood from leaking out. The clot also acts as a scaffolding where tissues and skin cells regrow, which is exactly the reason why wounds can heal very quickly.

Bad things can happen when a person does not have enough platelets needed to heal their wounds, especially in cases of severe injuries. This would lead to longer recovery time and loss of blood, which could result in death.

Because of this, 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 that daily demand for platelets is highly volatile and unpredictable. Because platelets donated from blood centers need to spend two days in testing before being certified as usable, hospitals have to predict their platelet usage two 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 excess platelets end up being wasted if they are not used before their expiration date.

Approximately 10.5% of 2 million platelet units are wasted annually, and this represents $80 million in healthcare costs! This large sum of money would definitely be better used to improve the quality and effectiveness of other areas in the healthcare industry, but instead, they are being wasted.

These problems make it very important for hospitals to accurately forecast platelet usage days in advance. An accurate prediction is vital to minimize wastage of platelets, while at the same time maintaining a buffer stock of platelet units as a safeguard measure.

Although there are prediction methods such as manual chart reviews, they are notoriously inaccurate and time-consuming. There have also been numerous studies that used patients’ data to predict platelet usage, but many of these had very limited patient populations and thus 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 by developing a mathematical model which would be able to forecast platelet usage three days in advance accurately. The result of the forecast -- the optimal quantity of platelet units that minimizes wastage -- would then be used to order platelets from blood donation centers and guide blood centers on how many donors to recruit, as well as prepare warnings in advance when the demand is predicted to surge.

Unlike previous studies, this time the research utilized large amounts of data collected over a very long period of time. They hypothesized that certain variables in patients’ data, such as platelet count, blood composition, and the number of patients in each hospital unit greatly influence the usage of platelet units. To test this hypothesis, 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.

They then trained the mathematical model using a software. Imagine a student who reads textbooks for more than 800 days non-stop. We quiz the student at the end of each day and give him his quiz score as well as the suggested answer. It is no doubt that he will learn from his mistakes and perform well for his final exam at the end of the period.

In the same way, our model is being fed data and learns the trend 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 at the end of 800 days was a model that can accurately predict platelet usage with little error.

At the beginning of the study, the hospital wasted 1 in 10 platelet units on average daily. Using prediction from this new model, only 1 in 30 platelet units are 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.

There were also interesting insights obtained from the resulting model. The model discovered that demand for platelets is strongly influenced by the number of patients in the Trauma and Surgery unit, which is likely because of the more severe nature of injuries in the unit. Surprisingly, usage is also lower on the weekends, presumably because there are fewer injuries during holidays.

Despite the potential cost savings, this model is still limited in use as the researchers at Stanford only used data from one hospital. Every hospital has different patient needs and populations. So, in order for the model to work in a different hospital, it needs to be trained again. But then again, the hospital needs to collect data and train the model for long periods of time so that the results are 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. of time so that the results are 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 neighbourhood hospital using this mathematical model.

(1013 words)

**References:** (APA citation style)

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