1.In patients having open heart surgery after they have left the operating room, what is the last measured red cell level (expressed as hemoglobin and typically abbreviated as Hb) before the first blood transfusion is given?

Chart, histogram

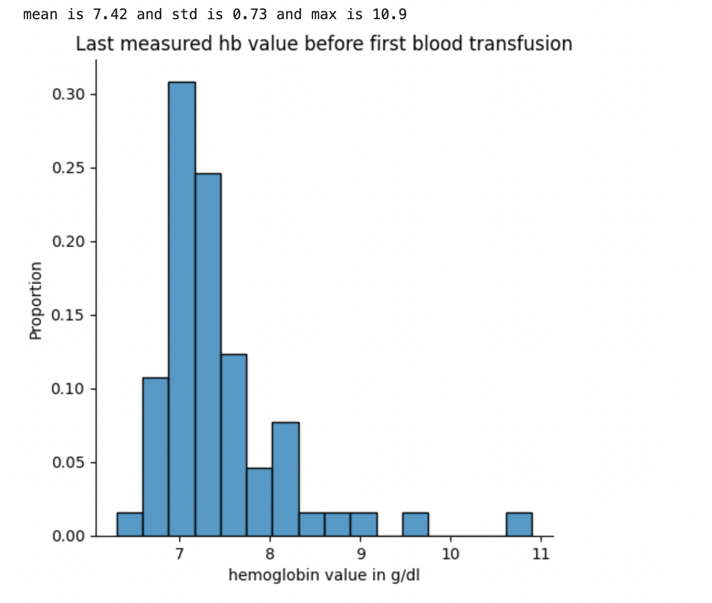
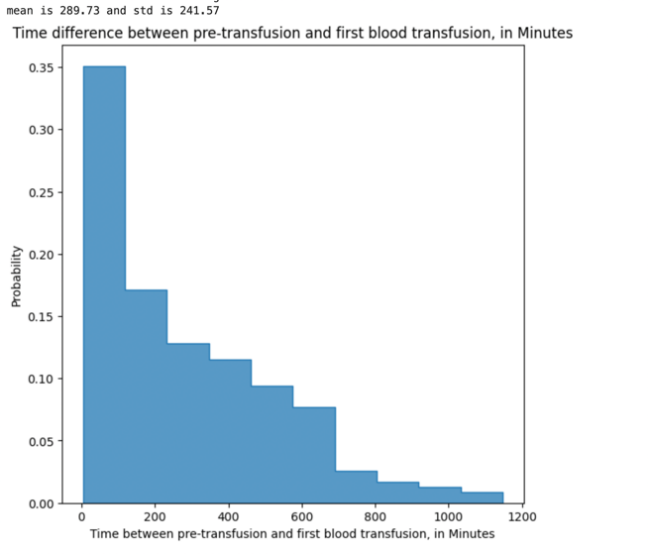
Description automatically generated

Figure 1 and 2 shows a distribution plot of the most recent hemoglobin value that was measured before the first blood transfusion. In terms of distribution MIMIC looks more like a normal distribution with mean 8.97 and standard deviation of 1.15. whereas RMH EPIC has a much more right skewed towards value between 7-8 with mean 7.42 and standard deviation of 0.73. Clinically the difference of hemoglobin value between the two dataset is very significant and p value produced from Mann Whitney U test is 10^-23. A possible conclusion that can be drawn from the plot is that doctors in RMH are more conservative when it comes to deciding whether a blood transfusion is necessary therefore doctor's threshold of hb value on when to allow a blood transfusion is much lower. However, the large difference might also be caused by other confounding factors, one of them been that the last measured hb value does not accurately represent the hb value during blood transfusion.In this study two possible factors were investigated that might contributes to this inaccuracy:

1. The time interval between hb measure and blood transfusion

This factor is quite obvious, the longer the patient has to wait for blood transfusion after the hb measurement, the higher the bias will be between the measured hb value and the actual blood transfusion thus in the clinical perspective, the difference of time between them should be minimized. In figure 3 and 4, a distribution plot of the time difference between MIMIC and EPIC is presented. When comparing both plots it’s apparent that most patients in RMH EPIC have a much shorter time interval. In terms of mean MIMIC has an average time interval of 289 minutes whereas RMH EPIC only has 130 minutes which is less than half of MIMIC. The statistical comparison using Mann Whitney U test again shows p value of 10^-10 which is very significant. Therefore, with such a big difference in time interval it’s very likely the RMH EPIC’s hb measurement has a better reflection on the actual hb value during blood transfusion compared to MIMIC.

Chart, histogram

Description automatically generated

1. Patient’s suffers from excessive bleeding after hb measure

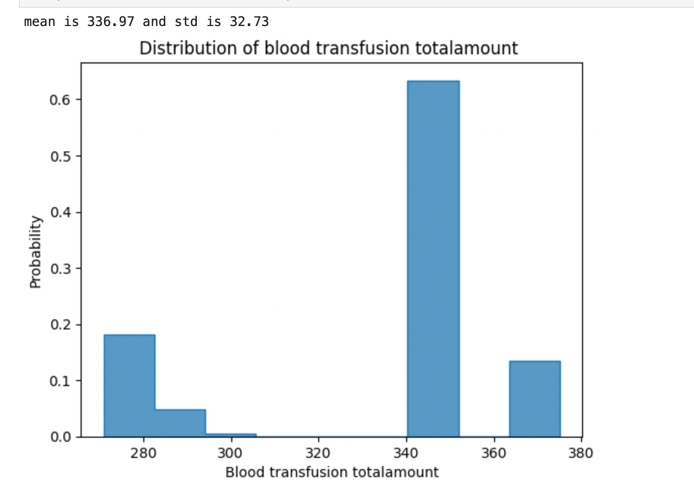
Another factor that can cause difference between hb value measured and actual hb value during blood transfusion would be excessive bleeding. This phenomenon can also explain why in the MIMIC dataset there are patients who receives receives blood transfusion even when their previous hb measurement is above 11.5(this is a common reference low point of hb value in RMH EPIC, patients should not receive blood transfusion if their hb level is higher than this value). Although we don’t have the data to pinpoint when and how much did each patient bleed, we can still test the hypothesis by plotting the correlation between the lastest hb measurement and the difference of the latest hb measurement made before and after the blood transfusion as shown on figure 5 and 6. If the value difference is negatively correlated with hb measurement value. Then it’s reasonable to assume blood transfusion given to a patient who has high hb measurement is due to the patient suffering from excessive bleeding because negative hb value difference indicates that the patient has bleed more than the blood he was given. When looking at both figures, we can see that the negative correlation exists in both hospitals. However,the magnitude of correlation is different. In MIMIC, the correlation is very clear and there are a lot of instances of high hb measurement with negative hb differences. On RMH EPIC the differences seem to be centralized at 0-2 with less proportion of patients that bleed more than the blood he was given. These results seem to suggest that excessive bleeding in RMH has been better handled than patients in MIMIC.

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

1. . In patients having open heart surgery,  after have left the operating room, when the first act of blood transfusion is performed, how many units of blood (also called red cells or packed red cells)  are given?

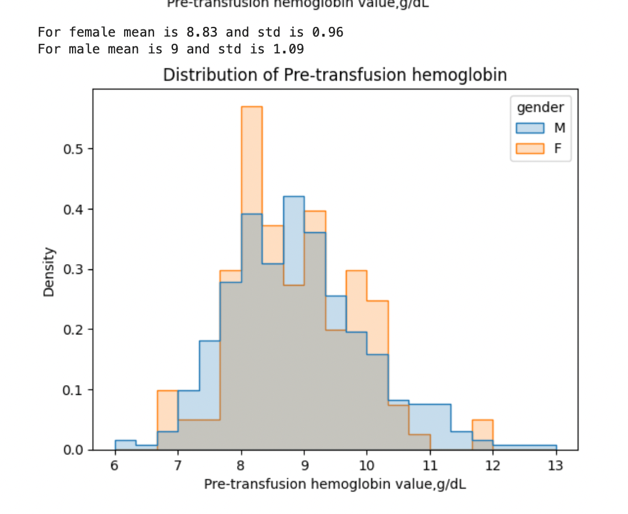
Chart, histogram

Description automatically generated

Figure 7 and 8 shows the distribution plot of the volume of blood that was given during the first blood transfusion in MIMIC and RMH EPIC dataset. The mean volume in MIMIC is around 337ml and in RMH EPIC is around 291ml and the Mann Whitney U test produced a p value of 10^-29 which suggests that the difference is very significant. However, one thing that should be noted judging by the data structure of the MIMIC’s plot is that the volume that has been recorded from the MIMIC’s side may not be precise. In figure 7 there are only 5 distinct values of volume represented as integer whereas in figure 8 the volume is a lot more variable which includes outliers within 0-100ml and volumes that is 500ml+. The volume itself is also more precise, up to one decimal place. This indicates that MIMIC’s data only tracks the volume of blood bag that was used for the patient rather than the actual amount of blood that was given to the patient. Additionally in MIMIC if two blood bags were used consecutively, they are counted as two separate instances, thus explaining why there is no volume that exceeds 375ml for the first blood transfusion in MIMIC.

1. . In patients having open heart surgery after they have left the operating room, what is the last Hb before a blood transfusion is given in men compared with women?

Chart

Description automatically generated Chart, histogram

Description automatically generated

Figure 9 and 10 shows a distribution plot of the most recent hemoglobin value that was measured before blood transfusion where the male and female density are drawn separately. Contrary to figure 1 and 2, all blood transfusions that happened after heart surgery are considered in these plots and with the addition of new data, only the RMH EPIC saw a very slight increase on the overall mean. While the overall distribution of hb value on the most recent hb measure before blood transfusion between MIMIC and RMH EPIC still shows a very significant difference(10^-44 which is even lower than the comparison between figure 1 and 2 due to the increase of degrees of freedom), gender does not seem to play a significant role on the two plot. In MIMIC, the p value of Mann Whitney U test when comparing the distribution of gender is 0.81 and in RMH EPIC the p value is 0.205.

4.In patients having open heart surgery, after they have left the operating room what is the difference between the first Hb after the blood was given and the last Hb before the blood was given?

Figure 11 and 12 shows a distribution plot of the hemoglobin value difference between the latest hb measurement taken before the blood transfusion and the first hb measurement taken after blood transfusion. The mean/standard deviation of hb difference in MIMIC is 0.99/1.3 and for RMH EPIC the mean/std is 0.81/1.13. According to Mann Whitney U test the difference between the dataset is not very significant with a p value of 0.08. However, from observing the plot it does suggest that RMH EPIC’s data has less extreme value on both ends which was previously shown on figure 5 and 6.

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

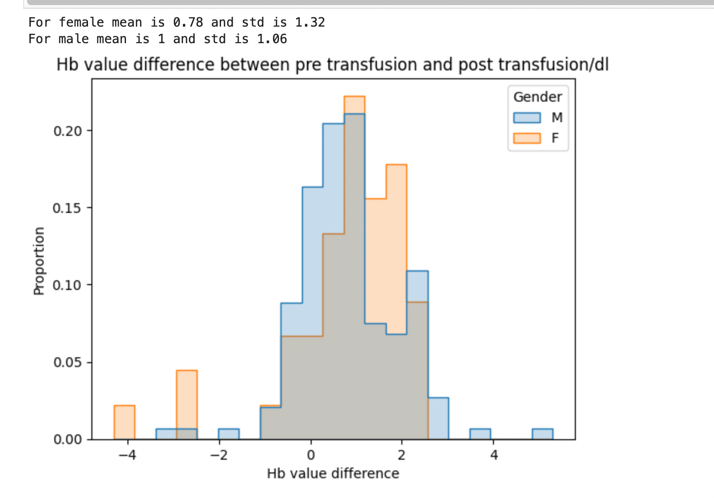
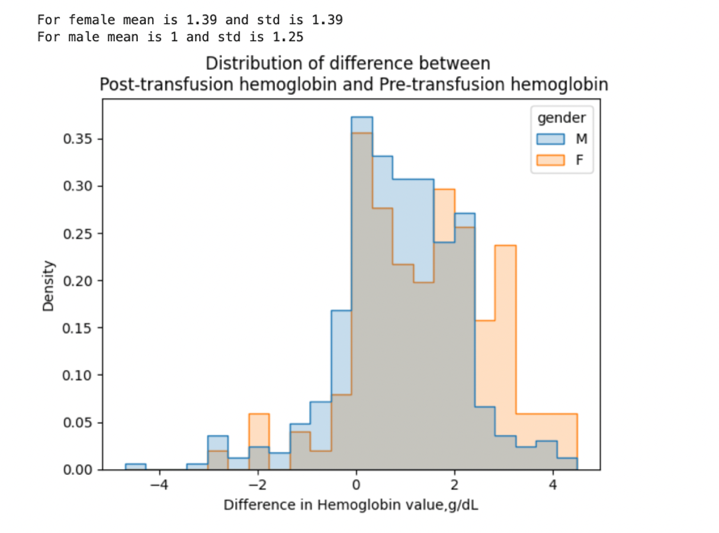
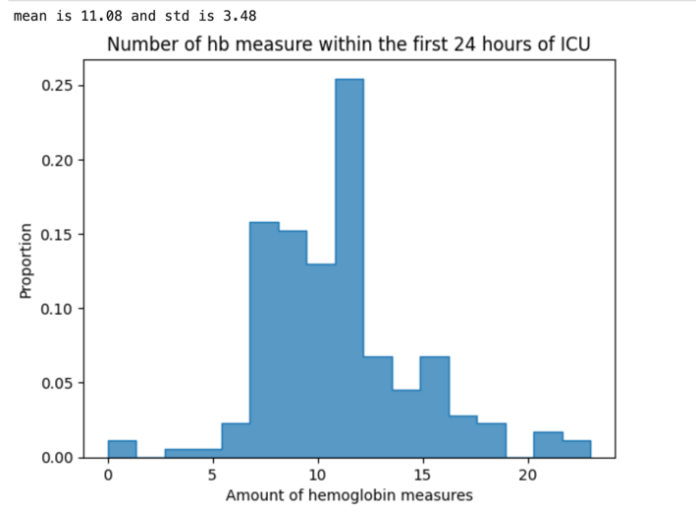
5. In patients having open heart surgery, after they have left the operating room, what is the difference between the first Hb after a single unit of blood wasgiven and last Hb before it was given in men compared with women? 

Figure 13and 14 shows a distribution plot of the hemoglobin value difference between the latest hb measurement taken before the blood transfusion and the first hb measurement taken after blood transfusion with male and female density plotted separately. In MIMIC, female has a higher mean of 0.39 between pre-transfusion hb value vs post-transfusion compared to male and the Mann Whitney U test produced a p value of 0.002 which is significant. In RMH EPIC, male has a higher mean of 0.22 between pre-transfusion hb value vs post-transfusion against male and the Mann Whitney U test produced p value of 0.38 which is not significant. In MIMIC, the results are inline with the expectation of medical experts. Since females have an innate lower concentration of hb value compared to male. By transfusing the same amount of blood to a female you will have a higher difference of hb value compared to male which is illustrated in figure 13 where there is a much higher proportion of females having a difference between 2-4 compared to male. In RMH EPIC, similar distribution were not shown which may be attributed to the scarcity of female’s sample(male has 4 times more transfusion samples compared to women)making the outliers of female sample have much higher weighting. Since there are no significant differences between the overall distribution without gender. It’s likely to be true that with more sample size the RMH EPIC’s female distribution will look closer to the MIMIC’s.

6. How many times was the patient’s hemoglobin measured during their first 24 hours of the intensive care?



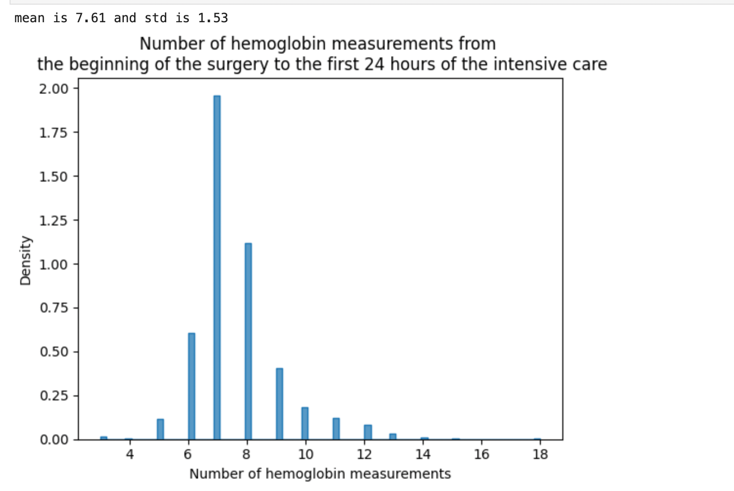


Figure 15 and 16 shows the total amount of hemoglobin measurements made during the first 24 hours of ICU in both hospitals and the difference are extremely large. In MIMIC the mean is 7.61 times and in EPIC it’s 11.08 which the Mann Whitney U test produced p value of 10^-64. Due to the lack of domain knowledge in the medicine field no comment can be made on these figues in terms of whether the number of hb measure in each hospital is appropiate nor the reason on why there are such a massive difference between the two hospitals. One speculation on the difference would be that MIMIC is understaffed which can heavily affect the frequency of hb test given to each patient. Understaffing has become a widespread problem in USA hospital as a mean of increase profitability and evidence of Beth Israel Deaconess Medical Center been understaffed can be found in forms of reviews[1] and journals articles[2]

[1]Weinberg, D. B., & Gordon, S. (2003). Not Enough Staff. In *Code Green: Money-Driven Hospitals and the Dismantling of Nursing* (pp. 137–159). Cornell University Press. http://www.jstor.org/stable/10.7591/j.ctt7zhdj.11

[2][Massachusetts National Guard deploys to help hospitals | The Journal Record](https://journalrecord.com/2021/12/28/massachusetts-national-guard-deploys-to-help-hospitals/)

Conclusion

In conclusion, this project has faced many challenges throughout the timeline. From the beginning it was the complex data structure of MIMIC and the lack of domain knowledge in the medical domain, to the conflict of interest on how clinical variability should be accessed from different medical experts, to the difficulty of acquiring RMH EPIC’s data due to all privacy issues. Nevertheless, this comparative analysis has covered all of the important questions that was requested by the medical experts and the main findings are as follows:

1. The overall variability of RMH EPIC’s cohort seems to be lower than MIMIC’s Cohort
2. The hb value measured before blood transfusion is significantly lower in the RMH EPIC’s cohort compare to MIMIC’s, and the difference applied equally to both genders
3. The time interval between hb value measured and blood transfusion is significantly lower in RMH EPIC’s cohort compare to MIMIC’s.
4. The Hb value measured for patient who stayed at ICU is a lot more frequent in RMH EPIC’s cohort compare to EP