EPIB 676: Final Report

Huzbah Jagirdar

Donor Return Dynamics: A Discrete Event Micro-Simulation Model

Background

Most medical care depends on a steady blood supply to meet urgent care needs in healthcare facilities. Blood collection sites often issue a hemoglobin deferral to donors when tested for low hemoglobin levels. It is shown that additional deferral decreases the donor's likelihood to return for a donation (Clement, 2021). Making an informed decision about potential changes to inter-donation interval (IDI) and deferral policies depends on how such polices impact long term donor return.

In a previous study, by increasing the IDI from 56 days to 84 days, for females, the deferral rate changes from 6.8% to 5.4% and the annual donation frequency decreased by 12% in females and 4% in males (Goldman, 2019). This highlights the importance of blood collection sites evaluating their inter-donation intervals and understanding the significance of encouraging donors to return promptly once eligible, while also minimizing the possibility of a hemoglobin deferral.

Method

A discrete event simulation model was implemented to assess three different inter-donation interval alternatives using comparative effectiveness and profit maximization.

Three three alternatives for inter-donation intervals being examined are:

1. The Standard: 56 days (8 weeks) interval for both males and females.
2. Alternative 1: 84 days (12 weeks) interval for females and 56 days (8 weeks) interval for males.
3. Alternative 2: 84 days (12 weeks) interval for both males and females.

These are being examined under the analytic framework of (i) comparative effectiveness which is being determined by the count of hemoglobin deferrals, count of successful donations and the deferral rate, and (ii) profit maximization, which is computed as the difference between revenue from a successful donation and cost of a hemoglobin deferral or successful donation.

Model Parameterization

The population sampled from are donors who donated with the South African Blood Donation Services (SANBS) between the years 2017 and 2022. Data from 2015 -2017 was used to computed donor history variables like red blood cells loss in the last 12 months, red blood cells lost in the last 24 months, days since last red blood cell loss, and days since last double red blood cell loss. We only used data from donors who donated or intended to donate a whole blood product or a red blood cell apheresis product for this model. Furthermore, donors who received permanent deferrals were excluded entirely. The population consisted of 3,658,560 successful donations and 439,099 deferrals of which 53% were hemoglobin deferrals.

Lasso Penalized Cox Proportional Hazards

Using the SANBS data, estimates were obtained for time to return after a successful donation or a hemoglobin deferral using a lasso penalized cox proportional hazards prediction model, adjusted for age at the time of visit, sex, race, donation product, donation at a fixed or mobile location, first time donation status and some donor history measure. The model was implemented in python using the sci-kit learn package and trained on ten percent of the data and produced the coefficients listed in table 1. The concordance index is 0.66, which indicates that the model differentiates between early events and later occurrences desirably. When scored using the test data, the concordance index is 0.67, an acceptable score.

````
```{python}
X_train3, X_test3 = train_test_split(X3, test_size=0.9, random_state=0)
cph = CoxPHFitter(penalizer=0.0001)
cph.fit(X_train3, duration_col = 'time_to_return', event_col = 'CENSORED')
cph.print_summary()

cph.score(X_test3)
```
````

TABLE 1

covariate	coef	exp(coef)	coef lower 95%	coef upper 95%
Visit_Age	0.00	1.00	0.00	0.01
first_time	-0.25	0.78	-0.26	-0.23
Fixed_mobile	0.20	1.23	0.20	0.21
cum_lifetime_donations	0.00	1.00	0.00	0.00
unit_rbc_loss	-0.27	0.77	-0.35	-0.18
rbc_loss_last_12_months	0.14	1.15	0.13	0.14
rbc_loss_last_24_months	0.06	1.06	0.06	0.06
days_since_last_rbc_loss	0.00	1.00	0.00	0.00
days_since_last_drbc_loss	0.00	1.00	-0.00	0.00
sex_F	-0.01	0.99	-0.44	0.42
sex_M	0.01	1.01	-0.42	0.44
donation_product_x_DEF PROD	-0.00	1.00	-0.86	0.85
donation_product_x_RBCAPH	0.03	1.03	-0.83	0.88
donation_product_x_WB	0.00	1.00	-0.85	0.85
race_African Black	0.03	1.03	-0.36	0.42
race_Asian	-0.13	0.88	-0.52	0.27
race_Mixed Race	0.03	1.03	-0.36	0.43
race_White	-0.00	1.00	-0.39	0.39
race_unknown	-0.05	0.95	-0.45	0.34
OUTCOME_TYPE_completed	0.00	1.00	-0.85	0.86
OUTCOME_TYPE_low hgb	-0.15	0.86	-1.00	0.69
OUTCOME_TYPE_other deferral	0.13	1.14	-0.71	0.97
first_time*fixed	0.10	1.11	0.07	0.13
first_time*hgb	0.09	1.10	0.02	0.16
hgb*fixed	0.02	1.02	-0.01	0.05

Logistic Regression

Using the SANBS data, estimates were obtained for the probability of a hemoglobin deferral using a logistic regression prediction model, adjusted for time to return and all the same covariates used in the cox proportional hazards model. The model was implemented in python using the sci-kit learn package and trained on thirty percent of the data using five fold cross validation and produced the coefficients listed in table 2. The model has a mean accuracy logistic regression classifier on test set score of 96%.

````
```{python}
#logistic regression to predict hgb deferral on follow up visit

#train model
````
````

````
X_train2, X_test2, y_train2, y_test2 = train_test_split(X2, Y2, test_size=0.7, random_stat

#print summary
model = sm.Logit(y_train2, X_train2)
result = model.fit()
summary=result.summary().tables[0]
summary.as_latex_tabular()

#train model
logreg = LogisticRegressionCV(cv=5, random_state=0)
logreg.fit(X_train2, y_train2)

#make predictions

y_pred = logreg.predict(X_test2)
print('Accuracy of logistic regression classifier on test set: {:.2f}'.format(logreg.score
```
````

TABLE 2

	coef	std err	z	P>|z|	[0.025	0.975]
time_to_return	-0.0038	5.25e-05	-72.733	0.000	-0.004	-0.004
Visit_Age	-0.0035	0.000	-9.149	0.000	-0.004	-0.003
first_time	-0.2094	0.020	-10.241	0.000	-0.249	-0.169
Fixed_mobile	-0.0601	0.011	-5.713	0.000	-0.081	-0.039
cum_lifetime_donations	-0.0005	0.000	-2.086	0.037	-0.001	-3.19e-05
unit_rbc_loss	-0.0105	0.225	-0.047	0.963	-0.452	0.431
rbc_loss_last_12_months	0.0883	0.008	11.524	0.000	0.073	0.103
rbc_loss_last_24_months	-0.0757	0.004	-17.980	0.000	-0.084	-0.067
days_since_last_rbc_loss	-0.0002	1.89e-05	-10.717	0.000	-0.000	-0.000
days_since_last_drbc_loss	0.0002	6.14e-05	2.475	0.013	3.16e-05	0.000
OUTCOME_TYPE_completed	-1.1829	nan	nan	nan	nan	nan
OUTCOME_TYPE_low hgb	0.3065	nan	nan	nan	nan	nan
OUTCOME_TYPE_other deferral	-1.0181	nan	nan	nan	nan	nan
sex_F	-0.2933	1.7e+05	-1.72e-06	1.000	-3.34e+05	3.34e+05
sex_M	-1.6012	1.74e+05	-9.18e-06	1.000	-3.42e+05	3.42e+05
donation_product_x_DEF PROD	-0.7116	nan	nan	nan	nan	nan
donation_product_x_RBCAPH	-0.9815	nan	nan	nan	nan	nan
donation_product_x_WB	-0.2014	nan	nan	nan	nan	nan
race_African Black	-0.0229	1.93e+04	-1.19e-06	1.000	-3.78e+04	3.78e+04
race_Asian	0.0065	1.93e+04	3.35e-07	1.000	-3.78e+04	3.78e+04
race_Mixed Race	-0.3043	1.93e+04	-1.58e-05	1.000	-3.78e+04	3.78e+04
race_White	-1.1475	1.93e+04	-5.95e-05	1.000	-3.78e+04	3.78e+04
race_unknown	-0.4263	1.93e+04	-2.21e-05	1.000	-3.78e+04	3.78e+04
first_time*fixed	-0.2368	0.049	-4.879	0.000	-0.332	-0.142
first_time*hgb	0.5775	0.060	9.672	0.000	0.460	0.695
hgb*fixed	0.0392	0.029	1.334	0.182	-0.018	0.097
````

Estimating Costs and Revenue

The costs and revenue are estimated from the perspective of the blood collection sites. This accounts for cost of staff, donor care and the actual test procedure, and the revenue gained from selling a blood unit. The cost estimates were obtained from Grieve R, et al (2018), who estimated the cost of a successful donation and a hemoglobin deferral to blood collection site in the United Kingdom. This value was then converted to South African Rand to obtain the value for SANBS. The revenue estimate was obtained from the SANBS price listing for a whole blood unit for sale to a public hospital.

TABLE 3

Costs/Revenue	Estimated Value
Cost of Successful Donation	R609.28
Cost of Hemoglobin Deferral	R211.83
Revenue from Successful Donation	R2291.30

Model Implementation

The micro-simulation model draws a sample of donors without replacement from the population who go to donate and are either deferred or complete a donation. When the donors enter the model their time to return to donation site is predicted by the cox proportional hazards model and if this time occurs after the completion of the inter-donation interval under each alternative, they are allowed to proceed. Next, using their predicted time to return, the probability of deferral is predicted by the logistic regression model, which decides whether the donor will be allowed to complete the donation. All the successful donations and deferrals are recorded until 10,000 donors have arrived after the prescribed inter-donation interval under each alternative. Finally, the count of successful donations and deferrals are used to compute the comparative effectiveness, deferral rate and estimated profit.

````
```{python}
# Discrete event simulation model

def simulate_donation(cph, logistic_model, idi, donor):
    donor=donor.to_frame()
    donor=donor.T

    donor_for_cph=donor.drop(labels=['time_to_return','CENSORED'], axis=1)
    # Use the Cox model to predict the time to return
    time_to_return = cph.predict_median(donor)

    # Simulate whether the donor returns based on the predicted time
````
````

```
    if np.random.uniform(0, 1) > 1 - np.exp(-time_to_return):
        return 0

    if (time_to_return < idi) and donor['OUTCOME_TYPE_completed'].any()==1:
        return 0

    donor_for_log=donor.drop(labels=['CENSORED'], axis=1)
    # Use the logistic model to predict the probability of hemoglobin deferral
    X = donor_for_log #change
    hemoglobin_prob = logistic_model.predict_proba(X.values)

    #print(hemoglobin_prob)

    # Simulate whether the donor is deferred based on the predicted probability
    if np.random.uniform(0, 1) < hemoglobin_prob[:,1]:
        return 1

    # If the donor is not deferred, record the donation and the inter-donation interval
    return 2
```

Uncertainty Analysis

Uncertainty analysis was added to the discrete event simulation by performing deterministic sensitivity analysis. Differently sampled input parameters were passed to the simulation model to observe how the output changes.

Results

TABLE 4

	Standard	Alternative 1	Alternative 2
Successful Donations Count	9417	9594	7335
Deferral Count	583	335	361
Deferral Rate	5.83%	3.35%	3.61%
Estimated Profit	R15716085.45	R16066336.83	R12870426.07

The results suggest that the standard 56-day IDI for both men and women has the highest deferral rate of 5.83% and alternative 1, 56 days for men and 84 days for women has the lowest deferral rate of 3.61%. Alternative 2, 84 days for both men and women, results in the fewest successful donations since the interval is too long and a lot of donors are turned away for returning early. Alternative 1, has the highest successful donations. It is also the most profitable alternative with an estimated profit of 16 million. Alternative 2 is the least profitable with an estimated profit of 12.8 million. The estimated profit of the standard IDI is 15.7 million.

The output of the deterministic sensitivity analysis is presented in TABLE 5. The results seem to be mostly in line with what was obtained earlier.

TABLE 5

	Standard	Alternative 1	Alternative 2
Successful Donations Count	9353	9540	7390
Deferral Count	647	367	362
Deferral Rate	6.47%	3.67%	3.63%
Estimated Profit	R15594879.05	R15968729.19	R12671425.34

Expansion Plan

If this initial analysis were to be developed into one for submission for peer-reviewed publication, these are aspects that would be added or improved:

1. Externally validate and cross validate cox proportional hazard model to evaluate and improve prediction accuracy. This was too computationally intensive for this project with the time constraints.
2. Improve logistic regression model. Ideally use a different machine learning classification method as to have a better precision and accuracy. It would also be beneficial to have a non-binary classifier to predict completed donation, hemoglobin deferral or other type of deferral.
3. Obtain better and more accurate cost estimates for donations and deferrals by consulting SANBS.
4. Add more complexity to the micro-simulation model by introducing the aspect of donation at a fixed or mobile site. It would also be beneficial to track one donor over the 7 year time-horizon to observe their donation patterns. To make the project less computationally intensive only 10,000 individual donations and the follow up visit was tracked. Another improvement would be to have a more robust probabilistic uncertainty analysis.
5. Performing a comparative analysis between a deferral patterns in the United States (Vitalant) and South Africa (SANBS) would add a novel aspect to the project for publication

Discussion

The results obtained from the simulation model are in consensus with the wider literature about tailoring inter-donation intervals. Having different intervals for men and women seems to be optimal in terms of profit as well as effectiveness. Although alternative 2, does result in reduced deferrals, it also reduces the number of successful donations in the same time-frame. Therefore, adopting a 56 day for men, 84 for women inter-donation interval is recommended to maximize profit and effectiveness.

References

1.Clement M, Shehu E, Chandler T. The impact of temporary deferrals on future blood donation behavior across the donor life cycle. Transfusion. 2021;61(6):1799–808.

2.Goldman M, Yi QL, Steed T, O'Brien SF. Changes in minimum hemoglobin and interdonation interval: impact on donor hemoglobin and donation frequency. *Transfusion*. 2019;59(5):1734-1741. doi:10.1111/trf.15155

3.Kim LG, Bolton T, Sweeting MJ, et al. Impact of a post-donation hemoglobin testing strategy on efficiency and safety of whole blood donation in England: A modeling study. *Transfusion*. 2023;63(3):541-551. doi:10.1111/trf.17277

4.Toivonen J, Koski Y, Turkulainen E, et al. Prediction and impact of personalized donation intervals. *Vox Sanguinis*. 2022;117(4):504-512. doi:10.1111/vox.13223

5.Grieve R, Willis S, Corte KD, et al. *Unit Costs and Assumptions (Trial-Based Analysis)*. NIHR Journals Library; 2018. Accessed April 20, 2023. https://www.ncbi.nlm.nih.gov/books/NBK535021/