ISYE 6501 : Project
4/21/2021

Table of Contents

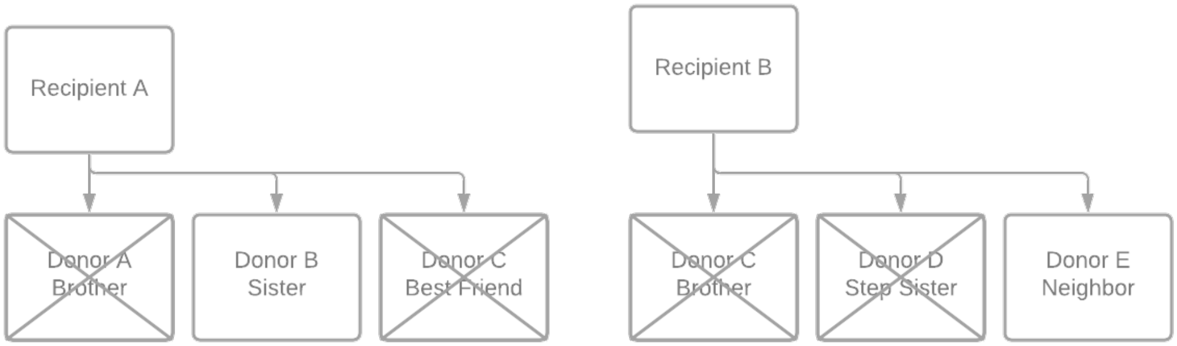
1 Project Overview
2 Model Set 1 - Probability
3 Model Set 2 - Prescriptive
4 Model Set 3 - Predictive

1. Project Overview

The project that I chose for ISYE 6501 is the Alliance for Paired Donation's (APD) study - Improving Kidney Donations to Save Lives (https://www.informs.org/Impact/O.R.-Analytics-Success-Stories/Improving-Kidney-Donations-to-Save-Lives).

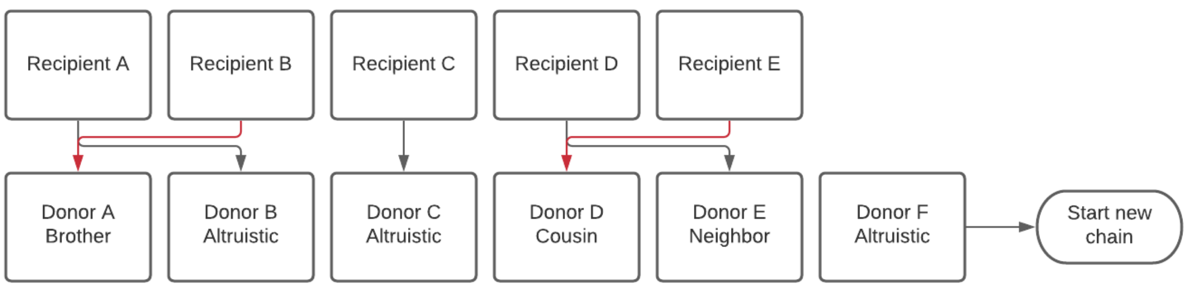
The Alliance for Paired Donation (APD) developed a process by which to match the maximum number of kidney donors with patient recipients. Past donor to recipient matches were primarily based on attempted matches with an individual who had a relationship with the patient recipient. There could be more than 1 potential donor with a relationship to the patient. If there were no compatible donors that matched by blood and tissue type, then the next option would be to wait for a deceased donor match. This creates a situation where the surgeries (or rapid delivery of a kidney) for transplants must take place simultaneously along with the matching of a patient. Living donor transplants generally have better outcomes for the recipient due factors such as kidney health and the ability to schedule the transplant based on recipient readiness (https://www.mhealth.org/childrens/blog/2017/july-2017/why-do-living-donor-kidney-transplants-offer-better-outcomes) .

Simultaneous pairing limits the number of available options for recipients, thereby potentially extending the amount of time that a patient is on dialysis. Highly sensitive donor receipients (such as those with blood type AB+) generally stay on the waiting list for a longer amount of time since potential matches are usually limited by waiting for a deceased donor.



Example of a Simultaneous Model

The abstract referenced the creation of non-simultaneous extended altruistic donor (NEAD) chains to significantly increase the pool of potential donors and possibly lessen the wait time for highly sensitive donor matches.



Example of Non simultaneous Model

I believe that an understanding on the feasibility of developing a repeatable NEAD solution to maximize the number of donor-recipient matches involve using the following 3 analytics model sequences:

- A model to evaluate the data features required to best maximize the number of donor-recipient pairings along with the probability of pairing failures.
- A prescriptive model to evaluate the optimal number of donor-receipient pairings to include in a non-simultaneous chain.
- A predictive model that evaluates the impact of changes in non-simultaneous (NEAD) chains.

2. **Model Set 1 – Probability**

The first step in evaluation is the determination of data that is available and useful. There are two primary factors in the determination of donor-recipient matching: blood type and tissue type matches. However, those are most likely not the only data points that would be of value in this situation.

Given: Collection of data features for the matching of donors to recipients (e.g., blood-type, tissue-type, previous recipient, pra (body's regularion of blood pressure, thirst and urine output), age, gender, weight, race, location)
Use: ***Lasso Regression***
To: Determine the data features that are the most relevant to determining the ability of a donor-recipient match and remove those data features that are not useful and will solely add complexities and processing time to modeling.

After determining the 1st round of data features most beneficial for matching donors to recipients, the next step will be to classify data to eliminate additional data to avoid overfitting.

Given: Output of data points from Lasso Regression
Use: ***K to the nearest neightbor KNN***
To: Classify individuals in 1 of the following 4 groups:

1) Good candidate for donation (no apparent underlying health conditions that make this individual a poor candidate for a kidney transplant)
2) Poor candidate for donation (apparent underlying health conditions that make this individual a poor candidate for akidney transplant)
3) Good kidney donor (no apparent underlying health conditions that make this individual a poor candidate to donate a kidney – e.g., they only have 1 kidney in use) and
4) Poor kidney donor (apparent underlying health conditions that make this individual a poor candidate to donate a kidney – e.g., they only have 1 kidney in use).

After reducing the data set to those features most relevant to the longer-term creation of a NEAD chain, I would then undertake the process of evaluating the probability of a successful transplant for the intented pool of donor-recipients based on age.

Given: Output of data points from the KNN evaluation
Use: ***Binomial Distribution***
To: Evaluate if age has an impact on the probability of success for any donor-recipient pairs.

According to historical data published by the Nephrology Dialysis Trsnaplantation periodical (https://academic.oup.com/ndt/article/27/4/1663/1833775) recipients of any age who receive a kidney from donors older than 65 years of age have a higher risk of rejection. The binomial distribution will be used to rank the data by probability of success. This is not to say that donor-recipient pairing where either individual is > 65 will automatically be excluded.

3. Model 2 - Prescriptive

With the features identified for the potential donors and recipients as well as the ranking for probability of success, data is now optimized to determine the number of chains that can be realistically created from the donor-recipient pool and when a new chain is required.

Given: Collection of donors (d) and recipients (r) from the first set of models
Use: ***Linear Programming Optimization***
To: Determine the maximum number of pairs that can be part of a NEAD chain and start new chains.

Linear Programming requires that 3 linear functions be defined:

1) Decision variables – Number of matched pairs based on blood-type and tissue-type matches for donors (d) – both altruistic and relational and recipients (r).
2) Objective function – determine the maximum number of pairs that can be created before a donor no longer has a match and a new chain must be formed.
3) Constraints – a new chain must be started when there are no longer recipients (r) available for a donor or a donor's match becomes unavailable (e.g., death or another match that is closer in location).

A variable bound must also be defined and in this case if there are no more donors available, d < 1 (obviously in whole integers since half a donor is not feasible), then the optimization problem cannot continue.

One additional level of optimization that could be done within each chain is to consider the location of donors and receipients in order to optimize the speed by which a kidney can be delivered to the recipient.

Given: A chain containing 1 or more matched pairs of donors and recipients and data on potential shipping methods (plane, helicopter, ambulance)
Use: ***Network Optimization Model***
To: Determine the fastest method of transporting the kidney from donor to recipient.

For this model assume that cost of shipping is not a factor since we are saving lives. Previous studies have shown that continuing on dialysis versus having a successful kidney transplant both shortens life, lessens quality and is more expensive over time (https://jamanetwork.com/journals/jama/article-abstract/1839743#:~:text=Transplantation%20and%20medical%20care%20costs,costs%20about%20%2444%2C000%20per%20year) .

4. **Model 3 - Predictive**

With relevant data factors having been selected and modeling for optimization in the matching of donors and recipients along with the best possible means of kidney transportation, I would now create a simulation model that can be run multiple times. The matched pair pool may not stay static. A donor may no longer be available due to health, location (moves to another country) or reconsiders involvement. A recipient may also no longer be available due to death or declining health no longer making the person a suitable candidate. Based on these uncertainties, I would run a simulation model to understand the impact of updating donors or recipients on the start of additional NEAD chains.

Given: Data for matched-pair NEAD chains, random selection of donors and/or recipients that no longer participate in the kidney transplant program.
Use: Stochastic simulation modeling
To: Determine the impact of losing matched pairs on reassignments within the chain and the need to start a new chain.

- The sample set of random values will be made up of donors and recipients who are part of a matching pair. All possible outcomes for 1 to n donors and/or recipients no longer being available for the kidney transplant program will be included.
- Probabilities will be assigned as follows:
    - 1- Donor is no longer available
    - 2- Recipient is no longer available
    - 3- Both matched-pair Donor and Recipient are no longer available.
- The events of interest are:
    - Any individual in a matched-pair NEAD chain is no longer available for the kidney transplant program.

The goal behind the simulation is to anticipate the likelihood of needing to have additional donors available that may be secondary matches for recipients.

Using the 3 model sets outlined in this project will lead to a larger pool of donors available for kidney transplant recipients and possibly shorten the time of transplants saving money and lives.