Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Econometrics and Operations Research

The impact of semi-directed donations on match quality in the Kidney Exchange Program.

Tessa De Weser (567581)

Supervisor:	Cremers, R.
Second assessor:	Wagelmans, APM
Date final version:	1st July 2024

Abstract

The Kidney Exchange Program (KEP), through which those experiencing kidney failure can match with other patients' donors, has been around for a little over two decades in the Netherlands. The topic has been highly researched in terms of the efficiency of its algorithm, allowing patient-donor pairs and altruists alike to participate in the Dutch National KEP. However, few have looked toward incentivising new altruistic or non-directed donors, who have the ability to help more than one patient with a single donation. These potential donors could be more inclined to participate when reassured that they help a young patient in particular. In this thesis, I aim to find out whether this phenomena, referred to as semi-directed donation, results in matches that are of lower quality than the average. I do so by studying the weights of the donations in an offline and online context. While the offline algorithm did not show significant results, the online model seems to indicate that semi-directed donors on average take part in higher quality matches. This hypothesis, in combination with the models provided in this thesis, is a starting point for the research that is left to be done on this type of altruism. The methods used in this thesis are applicable to an array of other potential incentives that could be used to increase the number of altruistic donors in the Kidney Exchange Program.

Contents

1 Introduction **1**

2 Literature **3**

3 Problem description **4**

4 Formulations **5**

4.1 Selective altruism models 5

4.1.1 Base model 7

4.1.2 Test model 7

4.2 Weights estimation 7

4.3 Selective altruism in an online algorithm 8

5 Data **9**

5.1 Offline algorithm 9

5.2 Online algorithm 10

6 Results **11**

6.1 Offline algorithm 12

6.2 Online algorithm 12

7 Conclusion **17**

A Code **IV**

B Replication of results **IV**

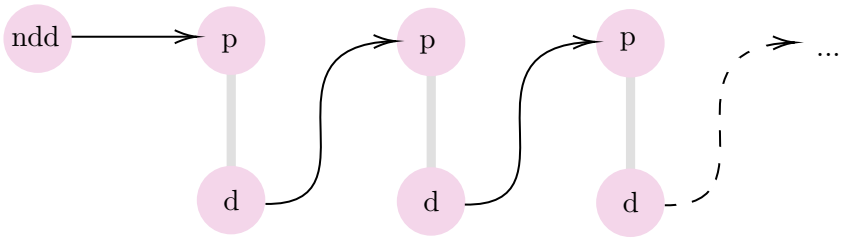
1 Introduction

According to the most recent report of the International Society of Nephrology (2023), kidney failure affects around 0.1 percent of the global population. The disease, characterised by an irreversible loss of kidney function, has become so prevalent that it is predicted to be the fifth leading cause of death worldwide by 2040 (Foreman et al., 2018). The preferred treatment for kidney failure is transplantation, but since the disease is so widespread the waiting list for deceased donor kidneys is much too long to cater to (Yi et al., 2021). Even if a patient is able to find a live donor, such as a friend or a relative, they are still not guaranteed to receive this kidney as a lot of donors are incompatible with their intended recipient. For those who do find a compatible donor, there is always the possibility that their body will reject the kidney after the operation.

To combat the shortage of kidneys, hospitals have started matching the patients of incompatible patient-donor pairs to other pairs' donors. This is called the Kidney Exchange Program (KEP). In its simplest form, it works by finding two patients with incompatible donors, who match with each other's donors, such that they both receive a kidney. This is called a 2-way or a 2-cycle exchange and it also works for more pairs than just two. Intuitively, a larger exchange would allow more people to receive a donation, but this does not translate into practise knowing the risks of kidney transplants. Last-minute tests may show that two people, who were assumed to be a match, are in fact incompatible (Constantino et al., 2013). To avoid transplanting only one person in a pair, hospitals try to perform the operations simultaneously and most of them will not have the capacity to facilitate large exchanges. Most KEP cycles include two pairs, after which 3-way exchanges are also fairly common. Both are usually implemented side by side, depending on the patient-donor pool (Kute et al., 2021).

Another form of exchange happens when a person, who is not linked with any patient, decides that they want to give one of their kidneys to a person in need. Since this donor does not expect a transplant in return, the patient receiving their kidney runs no risk of losing a paired donor if the operation fails. When such a non-directed donor (NDD) helps an incompatible pair, the donor of that patient can subsequently donate their kidney to the next pair, who then does the same. In the end, the last donor either donates their kidney to a patient on the deceased donor waiting list or they are part of a future chain (A. Roth et al., 2006). Thus, we are able to build a chain of patient-donor pairs around their donation, an example of which is shown in Figure 1.1. These transplants do not have to happen simultaneously, because a patient receives a kidney before their donor has to donate. Still, while in theory these types of k-exchanges can go on forever, in practise their length varies and there has been debate about whether or not longer chains result in more matches (John P. Dickerson, 2012).

Figure 1.1: Example of a chain started by a non-directed donor, 'ndd' , wherein 'p' stands for patient and 'd' for donor.



While NDD chains are an efficient way of helping patients who otherwise do not have a compatible donor, their implementation is directly limited by the number of altruistic donors. Many countries have strong legislative rules that apply specifically to kidney exchanges and, moreover, altruistic donation. Donors often go through a long process of psychological and medical tests. When they do check all the boxes, they can still be rejected because of quality issues (Ashlagi & Roth, 2021). Over the past two decades, several articles have appeared on the sorts of incentives that could bring more altruistic donors into the Kidney Exchange Program. However, these incentives have never been taken into account in Kidney Exchange models.

This thesis explores one of the proposed stimuli, in which an altruistic donor is assured that their donation goes toward a young patient. By restricting to whom it goes, this is no longer a non-directed but rather a semi-directed donation. I aim to answer the question: Do semi-directed donations result in matches of lower quality? Additionally, I seek to find out whether the overall quality of the other matches declines. The quality of a match is dependent on both the donor and patient's characteristic, to measure this I include a weights formula. The data used in this thesis are predominantly from Dutch sources.

By recreating the KEP algorithms put forward by Constantino et al. (2013) and adjusting them to fit the research question, I am able to form a Base and a Test model. The first acts as a measure of the number of patients that receive a kidney when there is a normal ratio of altruistic donors to pairs, including their respective weights. The second formulation tells us what would happen if we had more altruistic donors, but some of their donations are constricted by an age limit. An online algorithm is later applied to both models, in which a matching round is held in every quarter while people arrive to and leave the program.

This thesis provides important context to the philosophical debate surrounding semi-directed kidney donation. Methods used are simple enough to be applied to a wide array of other incentives, including but not limited to other donor preferences.

In Section 2 of this thesis I go through some of the current literature on incentivising donors, including the organisations that implement it and other factors that are taken into account when prioritising patients. Next, Section 3 contextualises the different elements to this problem, while Section 4 shows the models used to answer them. Section 5 contains the different sources from which data is taken in order relate the outcomes as closely as possible to the population of the Netherlands. Finally, Section 6 will go over the relevant results, while 7 summarises the findings.

2 Literature

The Kidney Exchange Program (KEP) was initially suggested in 1986, in response to the growing waiting lists for deceased donor kidneys (Rapaport, 1986). A first implementation of the program took place in South Korea in 1999, initially only allowing family of the patient to donate (Park et al., 1999). At first, kidney exchanges happened mostly within localised patient-donor pools. While there had been some earlier suggestions of models for a national KEP, Abraham et al. (2007) and A. E. Roth et al. (2007) proposed the first formulations that could handle the large patient-donor pools and match more than two pairs to one another. Their Edge and Cycle formulations were a breakthrough in terms of efficiency, solving the NP-hard problem to optimality for as much as ten thousand pairs at a time. The main idea of these models is that each incompatible pair is a vertex in a weighted, directed graph. Arcs then represent compatibilities between donors and patients of different pairs, such that cycles in the graph equate to exchange cycles. By enforcing a maximum size on those cycles, it became much easier to restrict the number of transplants to be performed simultaneously while also finding more and better matches.

Since 2007, both the Edge and the Cycle models have been used for many applications due to their general applicability. Constantino et al. (2013) has since suggested a more compact formulation, as well as an extension for the inclusion of altruistic donors. While there have been many improvements made to KEP algorithms in the past decade, the simplicity of these models in particular means that they are a great starting point for this thesis.

Altruistic donation has been a possibility in the Netherlands since the first matching round of the Dutch KEP in 2004 (Glorie et al., 2022). One non-directed donor can start a chain of non-simultaneous donations, offering opportunities for pairs that are otherwise unsuccessful in the KEP (Roodnat et al., 2010). Most recently, a study by Thomas et al. (2021) talks of luring in prospective altruists by, among others, directing their donation to a particular group of people. It hints that there is evidence to show that people may be more inclined to donate to children or underprivileged patients. A term used for this is 'moral particularism', which means that we prefer giving to those whom we feel connected with, such as a member of a community or network (Thomas et al., 2021). It could be argued that campaigns such as Dove's Kidney Donation for US Veterans is a form of the concept (Dove, 2024), or that smaller religious organisations have implemented similar tactics. However, these organisations connect new incompatible pairs, making them part of the direct donations.

A method of incentivising donors could be by allowing anonymity to be broken after the surgery. The United Kingdom is currently the only European country where this is possible. It has proven to be successful, with 92 percent of survey respondents saying that they broke anonimity and most being content with their decision (Pronk et al., 2022). Posner et al. (2013) has even suggested that the patient donate money to a non-profit of their donor's choosing, thereby walking the line of a monetary reward system.

A lot of these suggestions have never been included in the KEP because of obvious moral objections, but there has been one country to explicitly test the consequences. Israel openly engages in what is referred to as semi-directed donation with the help of the Matnat Chaim organisation. This organisation allows altruistic donors to set certain requirements for their

recipient, such as religious belief. Rather than using a linear programming model, the organisation uses direct matching of donors to patients who meet their criteria (Matnat Chaim, 2024). Between 2013 and 2020, the share of live kidney transplants in Israel that were realised through Matnat Chaim, has grown from 24 to 67 percent (Dienstag et al., 2024). Since its establishment however, the organisation has been the subject of criticism because of allowing donations to be determined by religious belief. Other research on the topic, by Dienstag et al. (2024), showed that donations to both Jewish and Arab recipients have increased.

Even if the results in Israel seem promising, they cannot be applied to the Netherlands. A near ninety percent of semi-directed donors in Israel identified as either observant or highly observant Jews in 2022, while at the time they represented only a fifth of Israeli society. In addition to that, while most Israeli donors wanted to meet the recipient of their kidney, a minority of Dutch non-directed donors feels the same way (Slaats et al., 2018). Outside of Israel, to the extent of this research, there have not been any instances of clear preferential treatment in altruistic kidney donation.

Currently, KEPs weights are based on more than simply the success of the transplant. Things such as a patient's time on dialysis or whether they have the same blood type as a potential donor can influence whether a match is chosen. In many countries, including the Netherlands, patients that are difficult to match will be prioritised when there is an available donor (Biró et al., 2021). Age and age difference are taken into account as well, with recent research showing that donations from older donors to older patients carry the highest risk (Hiramitsu et al., 2021). Moreover, a patient's Panel Reactive Antibody (PRA) level also contributes to the quality of their match. Being based on the percentage of antigens in the population for which a patient has developed antibodies, the PRA value is used to estimate the chance that a patient will reject their new kidney after the operation (Glorie, 2012). Regardless of the country's preferences however, almost all of them will first optimise the total number of matches before finding the best ones. This happens through optimising the model multiple times and forcing the new model to adhere to previous optima, a technique called hierarchical optimisation (Delorme et al., 2023).

As the patient-donor pool grows, it becomes increasingly important to balance quality and equity when matching patients. This thesis aims only to explore the possibility of including semi-directed donations in the KEP, without commenting on the moral implications of such a system. They are included in the altruist category for no other purpose than the clarity of this thesis and its models.

3 Problem description

In order to assess the effect of semi-directed donation on the quality of matches, two different scenarios were considered. The first scenario being a standard Kidney Exchange Program (KEP) with altruists. The second uses a similar structure but includes semi-directed donors (SDDs) into its mixed patient-donor pool. While these donors fall into the altruist category, they will only donate their kidney to a patient that is below a certain age. Both integer programming models used the principal of hierarchical optimization, where they first optimised the total number of matches and only then sought to maximise the sum of weights, which is the way that it is done in most KEPs (Biró et al., 2021). To test the effect of introducing SDDs over time, models

were studied in online algorithms with Poisson arrivals and departures. During a simulated time period of two years, both models received new admissions into their KEP at the beginning of each quarter while departures were scheduled after each matching round, at the end of the quarter.

To judge the quality of additional matches made with SDDs, the differences in their average weights were studied. This way, donations by semi-directed donors were compared to other, young patients who found a match in the KEP, as well as the overall averages in both previously mentioned scenarios. Weights are dependent on the age difference between the patient and donor, as well as the patients' crossmatch probability through their Panel Reactive Antibody level (Santos N., 2017). Both the online and offline algorithms ran for maximum cycle sizes equal to 3, 4 or 5, and for age limits of 25 and 35 years old.

4 Formulations

4.1 Selective altruism models

The models used in this thesis are based on the reduced Extended Edge formulation of Constantino et al. (2013). In this paper, the authors show how to adjust their model to accommodate non-directed donors (NDDs). This forms the foundation for the Base model and the Test model. The first one of which reflects a general KEP. The Test model includes additional altruists that can only donate to a subset of patients. Before going through the modified version of the integer programming problem, I first redefine the notation.

Define $V = \{1, \ldots, m+n\}$ as the set of all patient-donor pairs in the form of vertices. This includes m altruistic donors who are each matched with a dummy patient and n incompatible pairs. Since altruistic donors do not require a donation in return, their patient is set to be compatible to all paired donors $j \in \{m+1, \ldots, m+n\}$. Then, we define the directed, weighted graph G(V, A) and consider A the set of arcs. Vertices $i, j \in V$ are connected by an arc (i, j) if the donor in pair i is able to donate to the patient of pair j. Each arc has a weight w_{ij}, $(i, j) \in A$. Define k' and k as the maximum number of arcs in cycles with or without an altruistic donor, respectively. Then, take L copies of graph G, where $L = |V|$ is the upper bound on the number of cycles in the graph. Each copy $l \in L$ of the graph contains at most $max\{k', k\}$ arcs.

Every $l \in L$, if there is a cycle in copy l of the graph, represents the lowest index of any vertex used in that cycle. It does not include any other vertices whose index is lower than l. Hence, define d^l_{ij} as the shortest path distance between vertices i and j for $i, j \in \tilde{V}_l$ such that the path passes only through vertices of the set $\tilde{V}_l = \{i \in V : i \geq l\}$, (Constantino et al., 2013). This prompts the new set V^l for each vertex $l \in V$, which is the set of vertices who have a possibility of being in copy l of the graph. We define V^l differently for altruistic donors than we do for patient-donor pairs:

$$V^l = \{i \in V \mid i \geq l \text{ and } d_{li} + d_{il} \leq k'\} \quad \forall l \in \{1, \ldots, m'\}, \tag{1}$$

$$V^l = \{i \in V \mid i \geq l \text{ and } d_{li} + d_{il} \leq k\} \quad \forall l \in \{m'+1, \ldots, L\}. \tag{2}$$

Following this trend, define the set A^l to indicate the possible arcs in copy l of the graph as follows:

$$A^l = \{(i,j) \in A \mid i,j \in V^l \text{ and } d_{li} + 1 + d_{jl} \leq k'\} \quad \forall l \in \{1,\ldots,m'\}, \tag{3}$$
$$A^l = \{(i,j) \in A \mid i,j \in V^l \text{ and } d_{li} + 1 + d_{jl} \leq k\} \quad \forall l \in \{m'+1,\ldots,L\}. \tag{4}$$

Lastly, to further reduce the number of variables in this problem, l is only sampled from $\mathcal{L} \subseteq \{1,\ldots,L\}$, which is the set of indices l for which $\{V^l \setminus \{l\}\} \neq \emptyset$. Define the decision variables x_{ij}^l:

$$x_{ij}^l = \begin{cases} 1 & \text{if arc } (i,j) \text{ is selected to be in copy } l \text{ of the graph,} \\ 0 & \text{otherwise.} \end{cases} \qquad \forall l \in \mathcal{L}, (i,j) \in A^l.$$

The integer programming problem becomes as follows:

$$\text{maximise} \quad \sum_{l \in \mathcal{L}} \sum_{(i,j) \in A^l} w_{ij} x_{ij}^l, \tag{5}$$

$$\text{subject to} \quad \sum_{j:(j,i) \in A^l} x_{ji}^l = \sum_{j:(i,j) \in A^l} x_{ij}^l \qquad \forall i \in V^l, \forall l \in \mathcal{L}, \tag{6}$$

$$\sum_{l \in \mathcal{L}} \sum_{j:(i,j) \in A^l} x_{ij}^l \leq 1 \qquad \forall i \in \bigcup_{l \in \mathcal{L}} V^l, \tag{7}$$

$$\sum_{(i,j) \in A^l : i,j \in \{l\} \cup \{m+1,\ldots,L\}} x_{ij}^l \leq k' \qquad \forall l \in \mathcal{L} : 1 \leq l \leq m, \tag{8}$$

$$\sum_{(i,j) \in A^l : i,j \in \{m+1,\ldots,L\}} x_{ij}^l \leq k \qquad \forall l \in \mathcal{L} : m+1 \leq l \leq L, \tag{9}$$

$$\sum_{j:(i,j) \in A^l} x_{ij}^l \leq \sum_{j:(l,j) \in A^l} x_{lj}^l \qquad \forall i \in V^l, \forall l \in \mathcal{L}, \tag{10}$$

$$x_{ij}^l \in \{0,1\} \qquad \forall (i,j) \in A^l, \forall l \in \mathcal{L}. \tag{11}$$

The objective of this model maximises the weighted sum of all matches between donors and patients. When the weights are unitary this equates to maximising the total number of matches. Constraints (6) state that the number of kidneys given by a pair is equal to the number of kidneys that they receive, within each copy l of the graph. Constraints (7) make sure that each vertex belongs to at most one copy of the graph, thus, each copy will have a unique index l. The cardinality constraints in this model, in (8) and (9), each set the maximum number of arcs in every copy of the graph. The maximum allowed path length for cycles starting with an altruistic donor differs from those consisting of only incompatible pairs. What is important to note here, is that altruists will only ever appear in a copy of their own index. This is because they have the lowest indices of all vertices and it is not possible to make one chain that has two altruistic donors. Lastly, constraints (10) enforce that all copies of the graph indeed carry the index that is the lowest out of its vertices, while constraints (11) show that all of the decision variables are

binary.

4.1.1 Base model

Assume that the semi-directed donors (SDDs) have the lowest m' indices of all altruistic donors and thereby all vertices V. For the Base model, add the following constraint to (5) - (11):

$$\sum_{l \in \mathcal{L}: l \leq m'} \sum_{(i,j) \in A^l} x_{ij}^l = 0. \tag{12}$$

In the Base model, SDDs are not matched with any patients, thus the sum of their matches is equal to zero. The model will first optimise based on unitary weights w_{ij}, whereafter it is optimised based on the weights formula in Section 4.2.

4.1.2 Test model

In the Test model, there is only a subset of vertices $Q \in V$ who qualify for a donation from a semi-directed donor. These are the pairs whose patient falls below the age limit, thus $Q = \{i \in V : patientAge_i \leq ageLimit\}$. Therefore, for each $l \in \mathcal{L}$ where $l \leq m'$, copy l of the graph can only contain vertices that are in $Q \cap V^l$. To the model (5) - (11), add the following constraint:

$$\sum_{l \in \mathcal{L}: l \leq m'} \sum_{\forall j: (l,j) \in A^l, j \notin \mathcal{Q}} x_{lj}^l = 0. \tag{13}$$

This constraint ensures that no semi-directed donation can go to a person that is outside of the subset $\mathcal{Q}$. Again, this model first maximises the total number of matches before doing so with the weighted sum.

4.2 Weights estimation

Lack of academic uniformity and little data on the medical characteristics of participants in KEPs mean that there is no one way to estimate weights such that they accurately represent reality. Things like lifestyle, medical history and time on the waiting list no doubt play a role in the success of a candidate. However, in this thesis I focus only on two characteristics that have been proven to have an impact in the literature. That is, we look at a person's Panel Reactive Antibody (PRA) level and the age difference between a donor and their recipient. The specific role of the former was studied by Glorie (2012), and later adopted by Santos N. (2017) to generate the probability of a crossmatch for each patient. This probability c_{ij} , between the donor of pair i and the patient of pair j, is calculated using the cumulative distribution function of the standard normal distribution:

$$c_{ij} = \Phi(-1.5007 + 0.0170 \times \text{PRA}_j) \quad \forall j \in \{m+1, \ldots, L\} : (i,j) \in \bigcup_{l \in \mathcal{L}} A^l. \tag{14}$$

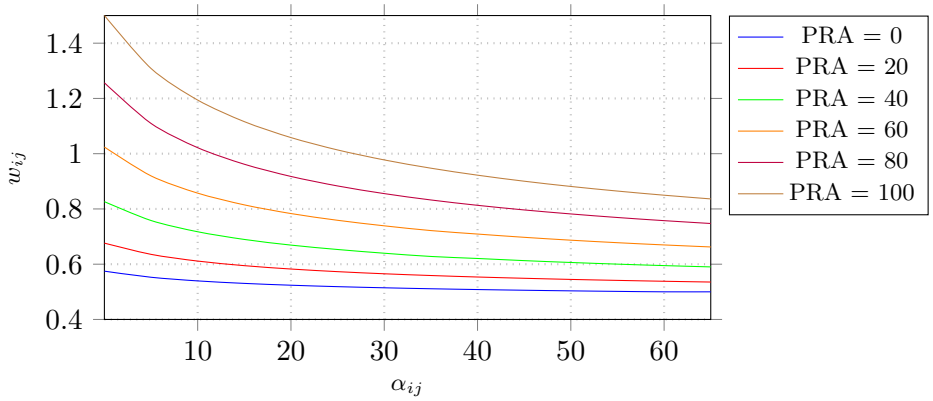
If there is a positive crossmatch between the blood of a donor and that of their intended recipient, then this means that the recipient has too many antibodies against the donor's cells and the operation can no longer take place. Age difference is also considered to have an effect on the

survival rate of the patient after the operation, with a number of countries having incorporated this into their weights already (Biró et al., 2021). For each match between a patient and a donor, where the patient is not an altruist's dummy patient, calculate the weight w_{ij} as follows:

$$w_{ij} = \frac{c_{ij}}{\sqrt{\alpha_{ij} + 10}} + \epsilon_{ij} \quad \forall j \in \{m+1, \ldots, L\} : (i,j) \in \bigcup_{l \in \mathcal{L}} A^l. \tag{15}$$

Where $\alpha_{ij} = |donorAge_i - patientAge_j|$ represents the absolute value of the age difference between the donor and patient. The error term is distributed as follows $\epsilon_{ij} \sim N(0, (\frac{\sigma_w}{h})^2)$, where σ_w is the standard deviation of a larges sample of weights before the error term is added, and h is a parameter to adjust the size of this standard deviation. As shown in Figure 4.1, setting the formula up in this manner ensures that if the age difference is very low, the match is assigned a high weight. Meanwhile, larger age differences will have less effect on the quality of the match. Furthermore, patients with higher PRA values are prioritised, since they have more antibodies and therefore a lower chance of being matched in the future. In Figure 4.1 as well as in the results of this thesis, weights have been standardised to fit between 0.5 and 1.5 before error terms.

Figure 4.1: Standardised weights versus age difference for various PRA values



4.3 Selective altruism in an online algorithm

In the online version of the models a simulation is run over a period of time, in which a new matching process takes place four times per year.

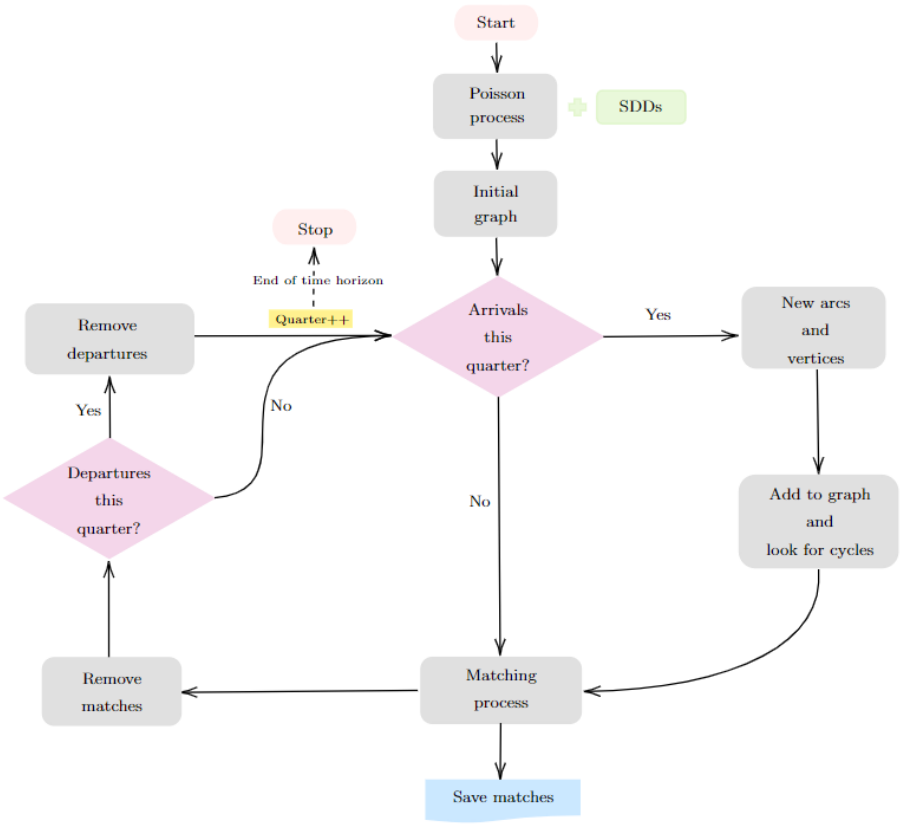
Before starting the optimisation process, arrival and departure times of pairs, altruistic donors and SDDs are generated using a Poisson process. For each new arrival, characteristics are determined and stored ahead of time, such that they remain the same in both the Base and Test model. All of the initial vertices and arcs are also added into sets and every arc is assigned a weight.

As seen in Figure 4.2 , at the beginning of each quarter, all of the vertices and arcs either from the initial graph or the end of the previous quarter are put into a new graph. Then, arrivals are incorporated and new weighted arcs are generated and added to the graph as well. Which new arcs to make is determined using the algorithm of Santos N. (2017). A patient and donor are compatible if the blood type of the donor in i and that of the patient in j is compatible, and

if a random number $r < PRA_j$, where $i \sim \mathcal{U}(1, 100)$. Then, finding the best cycles works with the Base and Test model to optimise the number of matches and total objective weights. After the optimisation, the weights of each individual match is stored, as well as their type. Then, the unused vertices and arcs of the graph are stored and any vertices or arcs pertaining to either the solution or the departures are removed. While the arrivals were added at the beginning of each period, departures are removed at the end. Finally, this process is repeated until there are no quarters left in the time horizon.

As seen in Figure 4.2, the only difference in the Test model compared to the Base model is whether semi-directed donors are included in the Poisson arrivals.

Figure 4.2: Online algorithm implementing the Base and Test models.



5 Data

In this paper, the aim was to find data concerning, or relating closely to, the Netherlands. All sources that were used are mentioned in the following subsections.

5.1 Offline algorithm

To determine how many altruistic donors to include in the patient-donor pool, I looked toward the Erasmus MC. They announced that, between January of 2018 and December of 2020, the Netherlands had seen an average of 32 unspecified donations per year (Joyce de Bruijn, 2022). Meanwhile, the average amount of people on the complete waiting list for a kidney in those same years (2018-2020) was 795 people (Nederlandse Transplantatie Stichting, 2024). Assuming

that the unspecified kidneys were distributed evenly over patients who are and are not part of the Kidney Exchange Program, this amounts to one unspecified kidney donation per 25 people. I also assumed that most altruistic donors, who were added to the KEP pool, found a match within one year. This way, I ended up with a simplistic ratio to generate three kinds of test instances, each consisting of 30, 50 or 80 incompatible pairs and 2, 4 or 6 altruistic donors, respectively. Half of the latter are considered to be semi-directed donors, which were randomly reassigned for every run. Note that, since there was no available data on the potential increase in donations if preferential treatment were to be implemented, this ratio is not the product of any calculations.

For the offline algorithm, test instances were generated based on the blood types of both the patient and donor, as well as the patient's PRA level from the United States National Kidney Registry. The altruistic donors were linked with a dummy patient that is compatible to every other paired donor. Per size, 10 different instances were provided by the Erasmus School of Economics.

To calculate the weight of a potential match, both the patient and the donor are assigned an age. For the patient's ages, I used data from the Dutch National Kidney Waiting List between 2018 and 2020 (Nederlandse Transplantatie Stichting, 2024). After adding the patients aged under 16 and above 85 to the nearest age category because of their low occurrence, I ended up with the following rounded numbers: 45% aged 16-55, 25% aged 56-64, 26% aged 65-74, 4% aged 75-85.

For the donor's age, this same distribution was used, under the assumption that the donors' ages are similar to those of the patients. While this is a grave assumption, it is known that around half of living donors are the husband or wife of their patient (A. E. Roth et al., 2007), and that most of the other paired donations are from friends or family. In the case of the donor being a friend or spouse, I assumed that their age is close to the patient's age. In the case of them being a relative, it is unclear what their relation is.

Lastly, out of 1 million estimates of the weights before error terms, the standard deviation was shown to be $\sigma_w = 0.2086$, this value was used in the final weights formulation of Section **??**. In addition, the parameter $h = 6$ was used, but can be replaced with other values according to the desired distribution. Furthermore, in section 6, the weights have been scaled to fit between 0.5 and 1.5 before error terms.

5.2 Online algorithm

The initial data set for the online algorithm remained the same as for the offline algorithm. Over a simulated two years, additional patients and donors were generated using a Poisson Process as well as the departure times of pairs in the KEP. The Poisson rates were based on data of the National Kidney Waiting List in the Netherlands (Nederlandse Transplantatie Stichting, 2024), which details the number of patients that were admitted to and let go from the waiting list between 2018 and 2020. After fitting these numbers to scale, they resulted in a rate of $\lambda_a = 0.521471 \cdot size$ for the arrivals and $\lambda_d = 0.467972 \cdot size$ for the departures, per year. Again, we use the ratio of 1 altruistic donor per 25 patient-donor pairs, meaning that the Poisson arrival rate of both NDDs and SDDs was $\lambda_a/25$.

There is little data available on the common blood types amongst patients signing up to the KEP. Therefore, it is hard to decipher the proportion of blood types that are the result of harder to match patients, who generally wait longer before their transplant. In this thesis, I chose to take the same distribution for candidate patients as for patients who are already on the waiting lists for deceased donor kidneys and the KEP. To simulate donors' blood types, a general estimate of the blood types in the Netherlands in 2020 was used (Haakman-Groot, 2020).

The allocation of an age to both the patient and donor in a new pair, happened in the same way as in subsection 5.1.

In order to generate a new patient's PRA level, I used a distribution based on patients from the Dutch National KEP between October 2003 and January 2011. Specifically, PRA levels with respect to the donor population. In Glorie (2012), these are categorized into three levels: 48% low (1-9), 35% medium (10-79) and 17% high (80-100). Each new patient was first assigned to either one of these levels with the respective probability, after which the exact PRA level was found through uniform sampling over the level's bounds.

To generate a new pair, a donor and patient were assigned characteristics and if their blood type was not compatible, they were immediately included in the KEP pool. If their blood types were compatible, then a random integer r between 0 and 100 would determine their overall compatibility. Only if the patients' PRA was above this random number ($r < PRA$), the pair could be included in the patient-donor pool. Pairs who did not meet this condition, whose PRA values were generally lower, were assumed to be compatible and were not included in this kidney exchange program. This method is based on the pool generation module in Santos N. (2017). While it does drive the overall PRA level distribution up, because more of the lower values are discarded, this effect is assumed to be minimal in an algorithm of this size. Over a sample of 1 million pairs, the average PRA levels were are around 13 values higher than when we did not discard compatible pairs and had a standard deviation that was 2 values larger.

6 Results

All of the following results were obtained with Gurobi 11.0.2 on IntelliJ 2023.2.4. with an Intel(R) Core(TM) i7-1065G7 processor at 1.30GHz and 16GB RAM.

For this thesis a large part of the results shown in Constantino et al. (2013) were replicated. As they are of no relevance to this thesis, I have included all outcomes and a short description in Appendix B.

As mentioned in Section 5, in both algorithms the Base and the Test model were run on instances of sizes 32, 54 and 86, which are referred to as sizes S, M and L from here on out. Each size ran with two different age limits for SDD patients, 25 and 35. Additionally, each combination of size and age limit was studied for three different values of k: 3, 4 and 5. Unless mentioned explicitly in the following subsections, the algorithms each ran for 10 different instances.

To further enhance the clarity of the tables in this section, weights in the Test algorithm will be shown as a the percentage increase of decrease in comparison to their Base counterparts. For SDDs this means they are compared to other matches of patients under the age limit made with the Base model.

6.1 Offline algorithm

For sizes M, the values given in Table 6.2 are the averages of 6 out of 10 instances for $k = 5$. The L data sets ran for all but 2 instances for $k = 5$, and all instances for other values of k. All other sizes and values of k were solved to optimality.

After a single matching round, it is shown in Table 6.2 that the average number of additional matches made in the Test algorithm compared to the Base algorithm ranges from 0 to 1.66 matches per instance. In most cases, the number of matches starting in an SDD is more than the increase in total matches. This suggests that these donors are used for more than extra transplants, but also replace old matches with higher weights. However, the effects of those weights are barely noticeable in the final outcome. This can be because of the small number of SDD donations or because the algorithm compensated other match qualities to use the SDD, dampening the positive effect on the final objective.

The average total weights per match in the Test algorithm, denoted by w_{total}, remain very similar to those using the Base model. For the smallest instances, in which only half will use the SDD when available, this is no surprise. The largest data sets' average Test objectives also remain the same if not slightly better for all three values of k.

Looking at the average number of semi-directed donors, especially in smaller patient-donor pools, not all of them ended up donating in this one matching round. This was likely because the 8 to 23% of patients they could have donated to were incompatible to them or offered no additional matches.

The representation of pairs whose patient is under the age limit in the Test algorithm was higher than in the Base model. In some cases, this lead to an over-representation of young patients compared to their presence in the KEP pool, denoted by $\mathcal{Q}$ in Table 6.2.

Lastly, p in Table 6.2 are the p-values of a Mann-Whitney U test that was performed on two disjoint groups of outcomes: semi-directed donations and all other donations to pairs in $\mathcal{Q}$. Aside from one outlier, for the smallest instance of $ageLimit = 35$ and $k = 3$, the weights of their matches did not differ significantly at the 5% level.

Overall, there was no indication that including semi-directed donors in an offline KEP causes any harm to the overall quality of the matches. Nor was there evidence to show that these kinds of matches will have significantly lower weights than other donations to people under the age limit. Even so, most of the SDDs had higher weights than their young counterparts, though this difference was not significant.

6.2 Online algorithm

The Online algorithm was solved to optimality for all instances and values of k. The averages of all instances are given in Table 6.3.

The introduction of semi-directed donors in the Test algorithm seems to have had a negative impact on the overall quality of matches in all cases. This could be the effect of matching more young patients, who have relatively lower weights. It could also be that the model compensates on quality in order to reach the most amount of transplants.

The Test model had a positive effect on young patients, in particular those who were matched with a semi-directed donor. There were few cases for which the average weight of the SDD

matches was lower than that of the young patients in the Base algorithm. In general, the positive effect much outweighed the negative. Especially when the age limit is set to 35 years old, the additional possibilities for semi-directed altruists cause their matches to be up to 25% higher than those in the Base algorithm.

In two of the largest cases, for patients up until age 25 and $k = 5$ and for the age limit being 35 and $k = 4$, the weights of SDD matches differ significantly from other young patients in the Test model, at a 5% level of significance. Because these were some of the largest groups of matches on which the Mann-Whitney U test was performed, it could be that other significant differences went unnoticed due to a small sample size.

The connected scatter plots in Figure 6.1 and Figure 6.2 show the matches with semi-directed donors compared to the outcomes for those same patients in the Base algorithm, for data set L and $k = 5$. If a red data point is not connected to a grey match, this means that the patient did not receive a kidney in the Base model. When a patient found a kidney donor in both algorithms, the quality of those donations was not too far apart for most. As shown in Table 6.1, with an age limit of 25, 2 of the matches including SDDs had a weight that was more than 1 standard deviation lower in the Test model compared to the Base model. This could be one of the reasons as to why this subgroup showed a significant difference in the Mann-Whitney U test. While no patients got a match that was more than a standard deviation higher, more of them did get a match with a positive difference larger than half of a standard deviation.

The time in which a patient received a transplant did differ for both models, with some patients in the under 35 instances having to wait an additional year or more depending on the algorithm they were in. In some cases, patients would receive their match sooner in the Test algorithm, but in others this was true for the Base algorithm. When patients get a later match in the Test model, it could be because the introduction of an SDD arrival made other, better matches possible. It is also a possibility that the solver simply found multiple solutions with the same objective.

Because new arrivals had different characteristics and thereby different weights, I will not compare the online results to those in Table 6.2. Furthermore, other than there being slightly more matches made with semi-directed donors when the age limit is 35 compared to 25, the data shows no clear patterns of overall improvement in the total number of matches. Most of the large increases relate directly to an increase in Base matches as well, suggesting that they could be caused by dissimilar characteristics in the KEP pool.

Overall, there is an indication that weights are on average higher for patients whose donor is an SDD than for other young patients. On the other hand, because the optimisation process first looks to maximise the total number of matches, the quality of other matches seemingly declines. A difference in both groups was proven to be significant only for the largest instances and certain values of k.

Figure 6.1: Weights of patients under 25 who matched with SDDs in the Test algorithm, red, compared to their weights in the Base algorithm, grey, and the quarter in which they were matched[1]. Results for instances of size L, with $k = 5$ and $ageLimit = 25$.

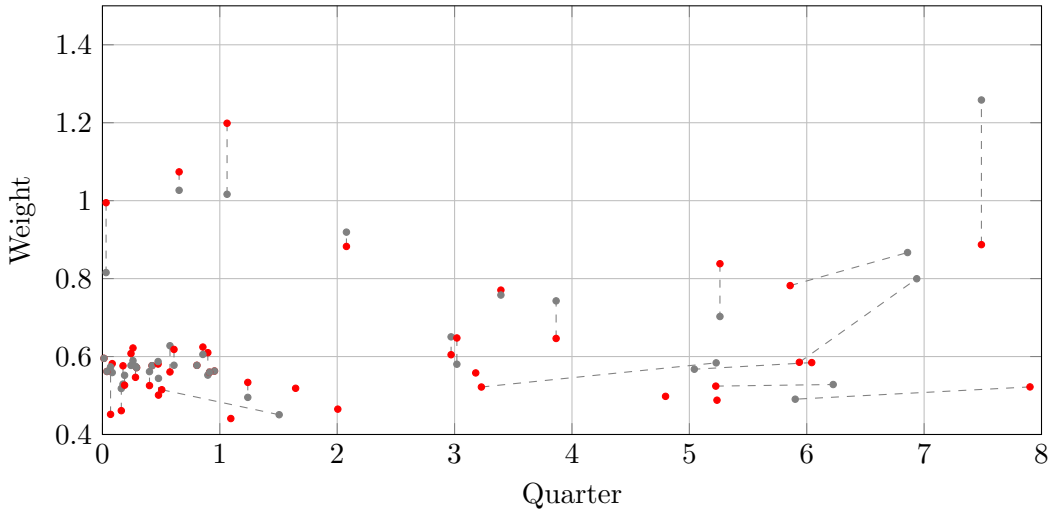


Figure 6.2: Weights of patients under 35 who matched with SDDs in the Test algorithm, red, compared to their weights in the Base algorithm, grey, and the quarter in which they were matched[1]. Results for instances of size L, with $k = 5$ and $ageLimit = 35$.

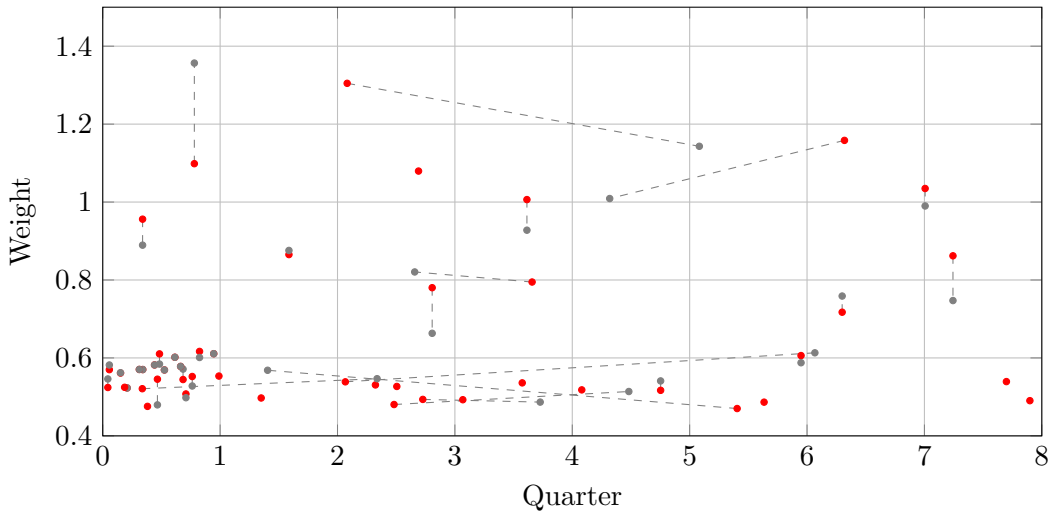


Table 6.1: Number of patients matched with SDDs in the Test algorithm, n_{sdd}, versus how many of them obtained a match in the Base algorithm, n_{base}. The last 4 columns reflect the number of patients whose SDD match was more than σ_w higher, more than σ_w lower, $\frac{\sigma_w}{2}$ higher or $\frac{\sigma_w}{2}$ lower than their Base match. Results for instances of size L, with $k = 5$ for age limits 25 and 35.

$ageLimit$	n_{sdd}	n_{base}	$> \sigma_w$	$< -\sigma_w$	$> \frac{\sigma_w}{2}$	$< -\frac{\sigma_w}{2}$
25	46	40	0	2	3	3
35	48	35	0	1	4	1

[1]Matches are spread out over the length of their quarter to increase visibility.

Table 6.2: Summary Statistics of the Base and Test offline algorithms

				Base				Test						
ageLimit	n	k	$\mathcal{Q}$	m_{total}	w_{total}	m_{age}	w_{age}	m_{total}	w_{total}	m_{age}	w_{age}	m_{sdd}	w_{sdd}	p
		3	10%	9.80	0.55	1.00	0.52	10.00	-0%	1.40	+0%	0.60	+3%	0.37
	S	4	13%	10.50	0.55	1.10	0.52	10.70	-0%	1.40	+0%	0.50	+2%	0.64
		5	13%	10.40	0.55	1.20	0.54	10.60	+0%	1.50	+2%	0.40	+4%	0.36
		3	12%	19.50	0.55	2.00	0.54	20.10	+0%	3.20	+1%	1.20	+2%	0.41
25	M	4	12%	21.00	0.55	2.40	0.54	21.40	-0%	3.40	-1%	1.30	-1%	0.99
		5	8%	17.00	0.56	1.00	0.54	17.00	-0%	2.00	+2%	1.00	+4%	0.63
		3	11%	34.50	0.55	4.50	0.54	35.40	+0%	5.70	+1%	2.30	-0%	0.52
	L	4	12%	35.70	0.56	4.50	0.55	36.40	+0%	6.20	+1%	2.30	+2%	0.30
		5	10%	34.50	0.57	3.50	0.56	34.63	+0%	4.63	+0%	1.75	+0%	0.75
		3	23%	9.70	0.54	3.00	0.54	10.30	-1%	3.50	-1%	0.60	-9%	0.00
	S	4	23%	10.60	0.54	2.10	0.53	11.00	+0%	2.30	+2%	0.50	+1%	0.88
		5	23%	11.00	0.55	1.90	0.55	11.00	+0%	2.40	-1%	0.60	-2%	0.42
		3	20%	19.90	0.55	3.60	0.55	20.50	+0%	4.60	-2%	1.10	-3%	0.40
35	M	4	22%	21.10	0.55	4.60	0.54	21.40	+1%	5.50	+1%	1.50	+1%	0.81
		5	22%	16.17	0.55	4.67	0.54	17.83	-0%	6.00	+0%	1.67	+1%	0.78
		3	20%	35.10	0.55	7.20	0.54	35.80	+0%	8.80	+1%	2.20	+3%	0.11
	L	4	19%	36.30	0.56	6.10	0.55	36.40	+0%	7.10	+1%	1.90	+1%	0.98
		5	22%	34.63	0.56	7.50	0.56	34.63	+0%	8.50	+1%	1.75	+0%	0.64

Notations: *ageLimit* denotes the maximum age of a patient to whom an SDD can donate, n the size of the data set, $\mathcal{Q}$ the size of the subgroup under the age limit compared to n (%), m_{total} the average number of matches, m_{age} those of patients in $\mathcal{Q}$, m_{sdd} those matches to an SDD, w_{total} the average weight of all matches (for Test this is compared to the w_{total} of Base), w_{age} and w_{sdd} are the weights compared to w_{age} Base (%). Lastly, p is the p-value of a Mann-Whitney U test on all matches in $\mathcal{Q}$ that were and were not by SDDs.

Table 6.3: Summary Statistics of the Base and Test online algorithms

			Base				Test						
ageLimit	n	k	m_{total}	w_{total}	m_{age}	w_{age}	m_{total}	w_{total}	m_{age}	w_{age}	m_{sdd}	w_{sdd}	p
25	S	3	30.90	0.60	1.20	0.54	31.70	-3%	2.10	+1%	1.00	+2%	0.67
		4	34.50	0.63	1.50	0.53	35.70	-4%	2.30	+1%	1.10	-1%	0.62
		5	33.10	0.60	1.40	0.55	35.00	-1%	2.40	+5%	1.10	+10%	0.34
	M	3	64.50	0.61	3.30	0.54	67.10	-2%	4.30	+1%	2.30	+2%	0.86
		4	62.90	0.61	4.80	0.55	68.40	-2%	6.30	+1%	3.20	+2%	0.34
		5	61.50	0.61	3.00	0.54	67.50	-1%	5.90	+3%	3.10	+4%	0.86
	L	3	93.10	0.64	5.00	0.54	93.30	-2%	6.90	+5%	3.30	+10%	0.08
		4	94.60	0.64	5.40	0.54	96.70	-1%	7.30	+5%	3.20	+11%	0.20
		5	100.70	0.63	5.70	0.54	103.50	-3%	9.20	+7%	4.60	+14%	0.00
35	S	3	36.90	0.60	1.70	0.52	37.90	-2%	2.20	+8%	0.80	+18%	0.06
		4	34.20	0.61	2.00	0.55	35.80	-0%	2.70	+1%	1.30	-0%	0.33
		5	35.20	0.60	2.30	0.53	35.80	-1%	3.00	+10%	1.10	+25%	0.27
	M	3	65.00	0.61	2.40	0.55	67.10	-2%	4.10	+2%	2.20	+6%	0.11
		4	70.90	0.62	3.90	0.54	72.90	-0%	5.50	+4%	2.70	+7%	0.60
		5	65.60	0.61	4.80	0.55	67.30	-2%	6.70	-0%	2.70	-1%	0.55
	L	3	93.20	0.64	4.50	0.54	98.10	-3%	6.50	+5%	3.60	+8%	0.70
		4	94.90	0.64	5.50	0.55	98.20	-2%	8.40	+8%	3.80	+19%	0.00
		5	93.00	0.64	3.80	0.54	97.40	-2%	8.10	+12%	4.80	+21%	0.06

Notations: *ageLimit* denotes the maximum age of a patient to whom an SDD can donate, n the size of the data set, m_{total} the average number of matches, m_{age} those of patients in $\mathcal{Q}$, m_{sdd} those matches to an SDD, w_{total} the average weight of all matches (for Test this is compared to the w_{total} of Base), w_{age} and w_{sdd} are the weights compared to w_{age} Base (%). Lastly, p is the p-value of a Mann-Whitney U test on all matches in $\mathcal{Q}$ that were and were not by SDDs.

7 Conclusion

In this thesis, I set out to study the effect of allowing certain altruistic donors in the Kidney Exchange Program (KEP) to donate only to young patients. More specifically, what would the inclusion of these kinds of donations mean for the quality of the matches, both directly and indirectly. While the debate surrounding semi-directed donors (SDDs) has been one of ethical concern, the aim of this thesis was to look into the efficiency of such a measure without making any further conclusions.

The methods used to investigate the quality of the matches were based on the compact Extended Edge formulation of Constantino et al. (2013). Their model formed the foundation for the Base and Test models, both of which were applied in an offline and online algorithm. Within the first context, donations by SDDs seemingly had a higher weight associated to them than other donations to patients under the age limit, which was either 25 or 35 years old. However, this effect was insignificant as per the Mann-Whitney U tests. The average weight over all matches remained nearly the same in every Test scenario, regardless of whether additional SDDs were a part of the KEP pool.

The results of the online algorithm were slightly more conclusive. In this context, average weights of donations made by SDDs were almost exclusively higher than similar patients' matches in the same KEP pool, and often much higher than those in the Base model. While most differences were insignificant, two of the largest instances did produce weights that differed from other young patients in their KEP pool at the 5% level. It is unclear if this was due to outliers or if semi-direct donors consistently provide higher quality matches. On the contrary, the average weights of matches in the Test algorithm were lower. This difference could potentially be attributed to the higher proportion of young matches in the outcome, who generally produce lower weights.

In general, there is an indication that donations by semi-directed donors produce matches of significantly different, and on average higher, quality. In the research question of this paper, it was assumed that these donors would produce lower weights, as they are limited in options. However, in the online algorithm, it seems that the benefit of being an altruistic donor outweighs the drawbacks of donating to only a subset of patients, especially when that subset is larger.

In future research, this hypothesis should be tested on instances with more patients and a known weight distribution. It should also implement common random numbers for the pool's characteristics of each test instance. Additionally, since the weights now rely on the age of patients, it could be interesting to use another formula or choose an unrelated characteristic on which to differentiate patients.

This thesis contributes to the literature by providing the models and algorithms to test differences in the KEP with and without semi-directed donors. Even if more conclusive proof would show that these SDDs indeed bring higher quality matches into the program, they also declined the quality of some of their patients' matches compared to the Base model. Thereby, one should be careful not to compensate the weight of other matches in order to include these SDDs and create more transplants.

References

Abraham, D. J., Blum, A. & Sandholm, T. (2007). Clearing algorithms for barter exchange markets: enabling nationwide kidney exchanges. In *Proceedings of the 8th acm conference on electronic commerce* (p. 295–304). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/1250910.1250954

Ashlagi, I. & Roth, A. E. (2021). Kidney exchange: An operations perspective. *Management Science*, *67*(9), 5455-5478. doi: 10.1287/mnsc.2020.3954

Biró, P., van de Klundert, J., Manlove, D., Pettersson, W., Andersson, T., Burnapp, L., ... Viana, A. (2021). Modelling and optimisation in european kidney exchange programmes. *European Journal of Operational Research*, *291*(2), 447-456. doi: https://doi.org/10.1016/j.ejor.2019.09.006

Constantino, M., Klimentova, X., Viana, A. & Rais, A. (2013). New insights on integer-programming models for the kidney exchange problem. *European Journal of Operational Research*, *231*(1), 57-68. doi: https://doi.org/10.1016/j.ejor.2013.05.025

Delorme, M., García, S., Gondzio, J., Kalcsics, J., Manlove, D. & Pettersson, W. (2023). New algorithms for hierarchical optimization in kidney exchange programs. *Operations Research*. doi: 10.1287/opre.2022.2374

Dienstag, P., Halaseh, R., Tzukert, K., Imam, A., Khalaileh, A. & Dienstag, A. (2024). An ethical analysis of conditional and partially directed live kidney donation among strangers - the israeli experience. *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons*. (Advance online publication) doi: 10.1016/j.ajt.2024.04.005

Dove. (2024). *Donor outreach for veterans.* Retrieved from `https://www.dovetransplant.org/about` (Accessed: 2024-06-13)

Foreman, K. J., Marquez, N., Dolgert, A., Fukutaki, K., Fullman, N., McGaughey, M., ... Murray, C. J. L. (2018). Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016-40 for 195 countries and territories. *Lancet (London, England)*, *392*(10159), 2052–2090. doi: 10.1016/S0140-6736(18)31694-5

Glorie, K. (2012). *Estimating the probability of positive crossmatch after negative virtual crossmatch* (Econometric Institute Research Papers No. EI 2012-25). Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute.

Glorie, K., Xiao, G. & van de Klundert, J. (2022). The health value of kidney exchange and altruistic donation. *Value in Health*, *25*(1), 84-90. doi: https://doi.org/10.1016/j.jval.2021.07.012

Haakman-Groot, M. (2020). Bloedgroep a, b, ab of o: waarom het ene bloed het andere niet is. *Nederlands Tijdschrift voor Hematologie*. Retrieved from `https://www.ntvh.nl/bloedgroep-a-b-ab-of-o-waarom-ene-bloed-het-andere-niet-is/`

Hiramitsu, T., Tomosugi, T., Futamura, K., Okada, M., Matsuoka, Y., Goto, N., ... Watarai, Y. (2021). Adult living-donor kidney transplantation, donor age, and donor–recipient age. *Kidney International Reports*, *6*(12), 3026-3034. doi: https://doi.org/10.1016/j.ekir.2021.10.002

International Society of Nephrology. (2023). *Global kidney health atlas.* Retrieved from `https://www.theisn.org/initiatives/global-kidney-health-atlas/` (Accessed: May 14, 2024)

John P. Dickerson, T. S., Ariel D. Procaccia. (2012). Optimizing kidney exchange with transplant chains: Theory and reality. *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, *11*.

Joyce de Bruijn. (2022). *Saskia gaf een nier aan een onbekende.* Retrieved from `https://amazingerasmusmc.nl/transplantatie/saskia-gaf-een-nier-aan-een-onbekende/` (accessed on 31/05/24)

Kute, V., Patel, H., Mishra, V. & Agarwal, S. (2021). Optimum chain length in kidney exchange program and the way ahead. *Indian J Nephrol*, *31*(2), 95–96. (PMID: 34267429; PMCID: PMC8240927) doi: 10.4103/ijn.IJN_376_19

Matnat Chaim. (2024). *Matnat chaim - volunteers for kidney transplantation.* `https://kilya.org.il/en/`. (Accessed: 2024-06-28)

Nederlandse Transplantatie Stichting. (2024). *Cijfers wachtlijst organen.* Retrieved from `https://www.transplantatiestichting.nl/medisch-professionals/cijfers-nts/wachtlijsten-organen` (accessed on 31/05/24)

Park, K., Moon, J. I., Kim, S. I. & Kim, Y. S. (1999). Exchange-donor program in kidney transplantation. *Official Journal of The Transplantation Society*, *67*(2), 336-338.

Posner, E., Choi, S. J. & Gulati, G. M. (2013). Altruism exchanges and the kidney shortage. (630).

Pronk, M. C., Burnapp, L., Reinders, M. E. J. & Massey, E. K. (2022). Relinquishing anonymity in living donor kidney transplantation: Lessons learned from the uk policy for anonymous donors. *Transplant International*, *35*. doi: 10.3389/ti.2022.10091

Rapaport, F. T. (1986). The case for a living emotionally related international kidney donor exchange registry. *Transplantation Proceedings*, *18*(3 SUPPL. 2), 5–9.

Roodnat, J. I., Zuidema, W., Van De Wetering, J., De Klerk, M., Erdman, R. A. M., Massey, E. K., ... Weimar, W. (2010). Altruistic donor triggered domino-paired kidney donation for unsuccessful couples from the kidney-exchange program. *American Journal of Transplantation*, *10*(4), 821–827. doi: 10.1111/j.1600-6143.2010.03034.x

Roth, A., Sönmez, T., Ünver, M., Delmonico, F. & Saidman, S. (2006). Utilizing list exchange and nondirected donation through 'chain' paired kidney donations. *American Journal of Transplantation*, *6*(11), 2694-2705. doi: https://doi.org/10.1111/j.1600-6143.2006.01515.x

Roth, A. E., Sönmez, T. & Ünver, M. U. (2007, June). Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences. *American Economic Review*, *97*(3), 828–851. doi: 10.1257/aer.97.3.828

Santos N., V. A. J. P., Tubertini P. (2017). Kidney exchange simulation and optimization. *Journal of the Operational Research Society*, *68*(12), 1521–1532. doi: 10.1057/s41274-016-0174-3

Slaats, D., Lennerling, A., Pronk, M. C., van der Pant, K. A., Dooper, I. M., Wierdsma, J. M., ... Massey, E. K. (2018). Donor and recipient perspectives on anonymity in kidney donation from live donors: A multicenter survey study. *American Journal of Kidney Diseases*, *71*(1), 52-64. doi: https://doi.org/10.1053/j.ajkd.2017.07.014

Thomas, R., Consolo, H. & Oniscu, G. (2021, Jul). Have we reached the limits in altruistic kidney donation? *Transplant International*, *34*(7), 1187-1197. (Epub 2021 Jun 19) doi: 10.1111/tri.13921

Yi, S. Y., Thomas, B. & Cooper, M. (2021, Sep). Updating deceased donor kidney allocation—what are the challenges? *Current Transplantation Reports*, *8*(3), 183–190. doi: 10.1007/s40472-021-00335-0

A Code

The code used to generate the results in this thesis are in a zip file. It contains four projects, one having been used to replicate the results of Constantino et al. (2013), another to generate large samples of weights and extract summary statistics while the last two run the offline and online algorithms of this thesis. As mentioned in the results section of this thesis, not all instances of the offline algorithm were able to run fully, while they were for the online algorithm. This means that the code for my online model is likely more efficient. Regardless, all other instances ran within at most a couple of minutes. Within the zip file I have also attached a document with steps showing how to recreate my results for each part of the thesis, but since each of the larger projects has an Interface class this will not take much time to understand. All of the projects used for this thesis were made with IntelliJ as a Maven project. If not familiar with Maven, it suffices to click the round arrow in the maven file of the project to download all of the dependencies that were used.

B Replication of results

To replicate both the Edge (E) formulation and the reduced Extended Edge (EE) formulation from Constantino et al. (2013), I used both blood type and density test instances. The latter are matrices that indicate the compatibility between donors and patients of different pairs, where the chance of a match between each is either 0.2, 0.5 or 0.7 . The other instances reflect a compatibility that is based on the blood types of both the patient and donor, as well as the patient's PRA level. These instances were provided by the Erasmus School of Economics (ESE). It is important to note that in either case, no donor is compatible to their own patient. Each data set consists of 50, 70, 100 or 200 incompatible patient-donor pairs and contains no altruistic donors.

While the model of this thesis does not seem more efficient, as it does not have fewer variables or constraints, it performs better for k = 4 and k = 5 in low density graphs. This is likely due to the improvements to this software over the last decade, and the use of Gurobi instead of CPLEX. Thereby, computation time was consistently lower for large instances, as shown in Table B. The notation used in both tables is the same as in Constantino et al. (2013):

- n is the number of nodes in the graph;
- k is the maximum length of the cycles;
- t_p is the average CPU time needed to solve for all paths in the graph, rounded to the nearest integer. In the EE model this time is negligible and thus not mentioned in Table B;
- T is the average CPU time it took Gurobi to solve the optimization problem, rounded to the nearest integer. The time limit for this was 1800 seconds;
- #opt is the number of instances that were solved to optimality within the given time limit. Each kind of data set, for each k, was run for 10 different instances. If there is no number given then this means that all were solved to optimality;
- gap is the average LP gap associated to the formulation. If UB is the upper bound provided by the linear relaxation of the problem, and Opt is the optimal value of the problem, then $gap = \frac{(UB-Opt)}{Opt} \cdot 100\%$;
- When a space is left empty, this means that the result was not relevant to the comparison. If a cell is filled with (-), this means that none of the instances produced results within 1800 seconds.

Table B.1: Results for small instances

n	k	E					EE			
		t_p/T	#opt	gap	#var	#con	T	gap	#var	#con
Blood type test instances										
50	3	10/91		0.0			0	0.0		
	4	28/447		0.0			0	0.0		
	5	54/1037	5	0.0			0	0.0		
	6	-		-			0	0.0		
Low density test instances										
50	3	0/2		0.0	2500	48651	0	0.0	6276	792
	4	0/1	8	0.0	2500	475449	1	0.0	6390	1810
	5	1/17	9	0.0	2500	$> 3 * 10^6$	1	0.0	6439	2315
	6	-		-	-	$> 3 * 10^6$	1	0.0	6572	2402
Medium density test instances										
50	3	0/2	9	0.0	2500	731982	4	0.0	17420	2024
	4	-		-	-	$> 3 * 10^6$	3	0.0	17273	2622
	5	-		-	-	$> 3 * 10^6$	4	0.0	17215	4101
	6	-		-	-	$> 3 * 10^6$	3	0.0	17579	2632
High density test instances										
50	3	0/8		0.0	2500	2073433	2	0.0	15345	2429
	4	-		-	-	$> 3 * 10^6$	4	0.0	26509	2640
	5	-		-	-	$> 3 * 10^6$	4	0.0	29440	3642
	6	-		-	-	$> 3 * 10^6$	3	0.0	29010	2641

Table B.2: Results for large instances

n	k	T	gap
Blood type test instances			
70	3	1	2.0
100		3	1.9
200		53	0.8
70	4	1	2.0
100		3	1.9
200		55	0.8
70	5	1	2.0
100		3	1.9
200		54	0.8
70	6	1	2.0
100		3	1.9
200		55	0.8
Low density test instances			
70	3	1	0.0
100		4	0.0
70	4	4	0.0
100		4	0.0
70	5	11	0.0
100		4	0.0
70	6	10	0.0
100		4	0.0
Medium density test instances			
70	3	7	0.0
100		24	0.0
70	4	16	0.0
100		24	0.0
70	5	13	0.0
100		24	0.0
70	6	11	0.0
100		24	0.0
High density test instances			
70	3	14	0.0
100		50	0.0
70	4	31	0.0
100		50	0.0
70	5	13	0.0
100		50	0.0
70	6	21	0.0
100		50	0.0