LUND UNIVERSITY

Priorities in Pairwise Kidney Exchanges - A Simulation Study on European Kidney Exchange Programs

Prioriteringar i parvisa njurbyten - En simuleringsstudie om europeiska njurbytesprogram

Master's Thesis, Faculty of Medicine

Samuel Lundgren (940817-5579)

Supervisors:

Tommy Andersson, *Professor at the Department of Economics, Lund University and Stockholm School of Economics*

Lars Wennberg, *Associate Professor and Senior Consultant at the Department of Transplantation Surgery, Karolinska University Hospital*

Assessor: Aboma Merdasa

Word count: 7984

Abstract

Background: For most patients with *end stage renal disease* (ESRD) kidney transplantation is the best treatment option. Best outcomes are achieved with living donation. To increase access to living donation for patients who are not compatible with their living donors, Kidney Exchange Programs (KEPs) have evolved. By exchanging donors, compatibility can be achieved thus enabling transplantation. In a so-called matching round many feasible, but conflicting, exchanges may be identified. Programs therefore implement rules for how to prioritize exchanges. The purpose of this study was to investigate how difference in priority rules between European KEPs affect which patients and donors are transplanted.

Methods: A side-to-side comparison was done by running KEP matching rounds with simulated pairs, based on real patient and donor data, and the priority rules from the British, Spanish, Dutch, and Scandinavian program. The simulations were limited to pairwise exchanges within the blood group barrier.

Results: With 150 pairs on average 51% were included in the matching round and of the matched recipients on average 72.6% were the same across programs. Significant differences were found in majority of the studied parameters. The British program seemed to match recipients with long waiting time, the Scandinavian those with higher cPRA (HLA immunization) and the Spanish younger recipients. These differences were clearer when increasing the patient pool.

Conclusions: The conclusion of the study was that many priority rules in the current practices likely cause systematic differences in which patients are transplanted and these differences are likely accentuated when the size of programs increase. As few similar studies have been conducted and little outcome data is available this shows importance of further research, especially as the future improvement of KEP effectiveness partly rely on increasing number of participants and international collaborations.

Populärvetenskaplig sammanfattning (Popular Science Abstract in Swedish)

Omkring 10% av Sveriges befolkning lider av kronisk njursjukdom. För de flesta patienter med långt gången, så kallad terminal njursjukdom är njurtransplantation den bästa behandlingen, både för patientens livskvalitet och ur ett hälsoekonomiskt perspektiv. Många patienter som har en villig donator passar inte med denna på grund av icke kompatibla blodgrupper eller för att de har antikroppar mot donatorns vävnadstyp. Njurbytesprogram adresserar detta problem genom att låta sådana inkompatibla par byta donatorer med varandra med målet att uppnå kompatibilitet inom de nya paren. Då detta kan öka tillgången till levande donation har flera sådana njurbytesprogram startats runt om i världen.

När man inom ett njurbytesprogram testar vilka par som potentiellt skulle kunna byta donatorer kan ofta flera möjliga byten identifieras. Då vissa av dessa byten kan stå i konflikt med varandra behöver man implementera regler för vilka byten som ska prioriteras. Dessa regler skiljer sig åt mellan njurbytesprogram. Målet med denna studie var att studera om, och vilka, av dessa regler som spelar roll för vilka patienter som blir transplanterade. Genom att simulera grupper av patienter och donatorer, baserat på riktig data, och därefter låta reglerna i det brittiska, spanska, skandinaviska och nederländska programmet välja vilka utbyten som skulle utföras kunde detta studeras.

Det som kunde ses var att andelen av matchade patienter som var den samma mellan alla program var i snitt ca 73% (något beroende av antal par). De skillnader som sågs mellan program var i hur länge patienter väntat på transplantation, ålder, åldersskillnad mellan donator och patient, hur sensitiserade patienterna var till andra vävnadstyper och grad av skillnad mellan vävnadstyp hos patient och donator. Dessa skillnader blev mer systematiska ju fler som ingick i dessa program.

Denna studie ämnar inte porträttera exakt hur njurbytesprogrammen presterar i sin naturliga miljö, men resultaten ger anledning att tro att det finns skillnader i utfall beroende på vilka prioriterinsregler de har. Då det finns få liknande studier och den tillgängliga utfallsdata är begränsad påvisar detta behov av vidare forskning och att öka standardiserad rapportering för att kunna jämföra program med varandra.

Table of contents

Background

Kidney transplantation and kidney exchange

For most patients with *end stage renal disease* (ESRD), kidney transplantation is the best treatment option. With modern immunosuppressive treatment and surgical protocols kidney transplantation prolongs survival, increase life quality, and decrease costs compared to dialysis (1, 2). However, access to transplantation is limited. In Scandinavia, the number of patients waiting for kidney transplantation has increased by 27% 2011-2021, despite increased number of kidney transplants. Eurotransplant, an organization coordinating organ transplantation activities among several European countries, also report increasing waiting lists for the same time period (3). Kidney transplantation is possible from both living and deceased donors, however best outcomes are achieved with living donors. Whilst transplantation from deceased donors has increased in Scandinavia the last 10 years, living donation has remained approximately constant, as can be seen in Figure 1 (4, 5). In the US, total kidney transplantations are up 40% 2008-2019, whilst transplantations from living donors have remained stagnated during the same period, despite a slight increase 2018 and 2019.

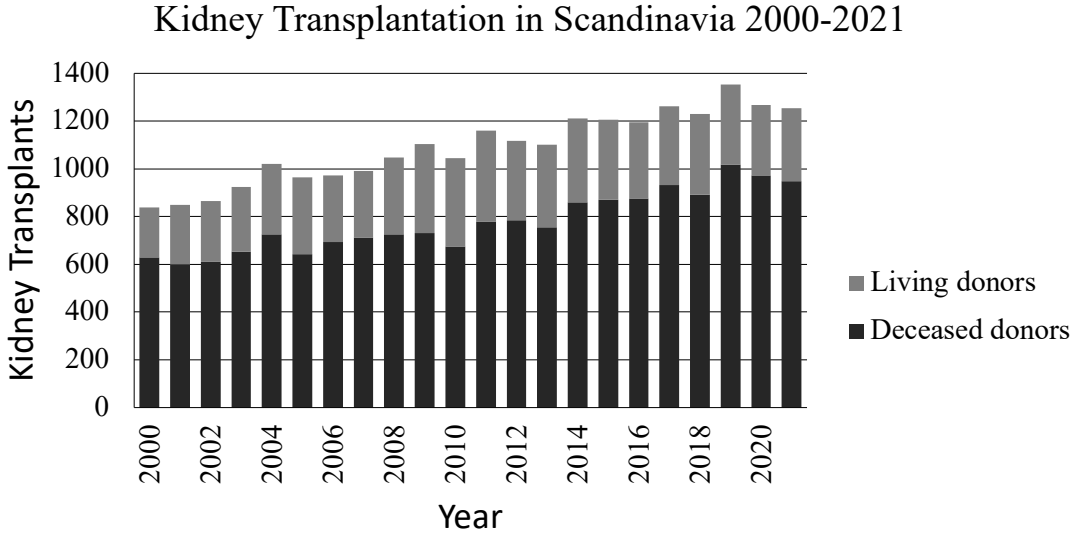


Figure 1: Number of kidney transplantations in Scandinavia 2000-2021 per donor type. Data from Scandiatransplant

Some of the main obstacles for living donor kidney transplantation are AB0 blood group incompatibility (AB0i), and *human leucocyte antigen* (HLA) incompatibility (HLAi) between the donor and recipient. Incompatibility is defined as the recipient

having unacceptably high titers of antibodies against donor specific AB0 blood-group antigens or HLA-antigens. Treatment modalities aiming at reducing such donor-specific antibodies, thereby enabling AB0i or HLAi transplantation, are referred to as desensitization protocols. Modern desensitization protocols for AB0i kidney transplantation have yielded long term outcomes comparable to those of AB0-compatible transplantation (although at a higher cost and level of immunosuppression). Desensitization protocols for HLAi kidney transplantation have not been equally successful and HLAi transplantation is associated with higher rejection rates and worse long-term outcomes compared to the compatible case (6, 7).

In the setting of organ allocation and transplantation, HLA-compatibility is evaluated by so called *virtual crossmatching* where the recipient's anti-HLA antibody repertoire is compared to the donor HLA-type. A positive crossmatch indicates that the patients have *donor specific antibodies* (DSAs) towards donor HLA-antigen and subsequently an increased risk for postoperative graft rejection. Such antibodies may develop after exposure to foreign HLA-antigens from previous transplantation, blood transfusion, or pregnancy. A *virtual crossmatch* is always confirmed with a *laboratory crossmatch* before surgery. The risk of a positive crossmatch against a random organ donor can be estimated by measuring *panel reactive antibodies* (PRA) or *calculated panel reactive antibodies* (cPRA). In commercially available PRA assays, patient serum (containing possible anti-HLA antibodies) is applied to a panel of HLA-antigens from donors considered representative for the population of potential donors. However, PRA does not directly consider the frequency of specific HLA antigens in the donor population. A cPRA comparing recipient anti-HLA antibodies with the frequency of HLA-antigens in different donor databases may be more accurate. A cPRA value of 45% indicates that the recipient has antibodies against 45% of the donor population. Hence, high cPRA/PRA suggests that identifying an immunologically suitable living or deceased donor will be challenging. Patients with high PRA/cPRA are often referred to as sensitized. HLA compatibility is to be separated from HLA-matching, which means evaluating similarities in recipient and donor tissue types, not reflecting sensitization level (8). Although a close HLA-match can improve outcomes, it does not play the pivotal role of HLA-compatibility.

With the aim of enabling transplantation also to patients with medically suitable but immunologically incompatible living donors the first European *kidney exchange program* (KEP) was initiated in the Netherlands in 2004 (9). In KEPs, the fundamental

idea is to pool together incompatible donor-recipient pairs to generate compatible pairs by exchanging donors in-between the included pairs. In its simplest form, this is accomplished by performing a *2*-way exchange, in which a donor from an incompatible pair donates to a recipient in another incompatible pair and vice-versa. In this way none of the matched recipients need to wait for a deceased donor, thus saving kidneys for ESRD-patients without a suitable living donor. In total, more kidneys become available for transplantation and more ESRD-patients can be helped.

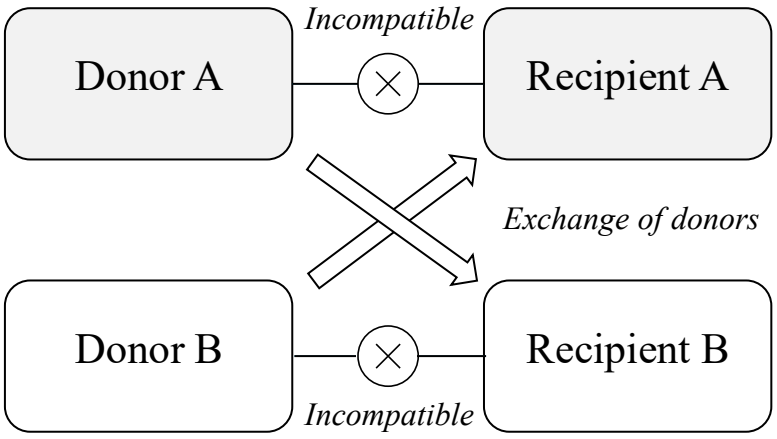


Figure 2: Schematic representation of a pairwise kidney exchange. Donor A is not compatible with recipient A, and same situation in pair B. The recipients are compatible with the other recipient's donor, thus motivating a donor exchange

Since 2004 a variety of additional kidney exchange programs with different arrangements has been established in Europe. The largest KEPs in Europe 2015, in terms of donor-recipient pool size, were the programs in United Kingdom, Netherlands, and Spain (9). From April 2018 to March 2019, 97 individuals were transplanted through the program in the UK. During 2019, 14 patients were transplanted in Holland and in Spain 33 (10-12). In Scandinavia, in the Scandiatransplant Kidney Paired Exchange Program (STEP), 49 transplants were performed from 2019 to 2021 (13).

In a KEP, it is necessary to define feasible exchanges. It is also necessary to prioritize which, of sometime several feasible exchanges that eventually should be performed, as some of these might be conflicting. A common priority is to perform as many transplantations as possible. Other examples are to prioritize based on sensitization level (cPRA), dialysis vintage, time on waiting list and time in the exchange program.

Although many KEPs strive to prioritize sensitized patients, these still tend to accumulate. In the Spanish program, the proportion of highly sensitized patients in the donor-recipient pool increased from 20% to 50% between 2010 and 2015 (9). Another such group is blood 0 patients. As for blood transfusions, kidney transplantation is normally performed from blood-group compatible donors. Consequently, type 0 imbalance is another major limiting factor, as type 0 patients have anti-A/B antibodies against all except type 0 donors. Although AB0i transplantation tend to increase match rate, even where AB0i-transplantations are allowed, some disadvantage for blood group 0 patients may remain (14). Another parameter to consider is age, as it is shown that donor age affect transplantation outcomes (15). Therefore, age-matching is a well-established principle in organ allocation aiming at avoiding large age-differences between donors and recipients. For kidney transplant patients, dialysis vintage is correlated to transplantation outcomes. Ideally, time in dialysis treatment should be minimized and transplantation should preferably be performed pre-emptive, i.e., before starting dialysis. How priorities are implemented in the programs vary, which we will be seen later.

Kidney exchange and priority rules in previous research

Pétér Biró and number of researchers wrote an overview of KEPs in Europe in 2019 (9). They identified three main areas of opportunity: extending national pools, allowing new modalities in exchange, and increasing international collaboration. Increasing pool size is one of the most fundamental principles in KEPs as this will increase the potential gains (possible matching percentage). Larger pool sizes and international collaborations comes with legal, logistical, medical, and ethical challenges putting increase pressure on KEP design. Larger pool size will also put high pressure on priority rules as the number of feasible exchanges and possible solutions will increase.

There are several articles discussing priority issues and new ways to prioritize. An evaluation of one of the American kidney exchange programs, APD, identified accumulation of highly sensitized patients as the main issue. It also showed that non-sensitized (low cPRA) blood group A received grafts within 73 days, whereas non-sensitized blood group 0 waited on average 335 (16). Wenhao Liu and Marc L. Melcher showed in another article that prioritizing hard-to-match patients can, over time, improve waiting time in KEPs even for easy-to-match patients (17). They also showed that using metrics considering not only how many donors a patient is compatible with (as commonly done in many KEPs), but how many patients that patient's donor is

compatible with can improve the overall waiting time in the pool. In another article, Tiago Monteiro and colleagues showed that using matching time as highest ranking objective can over time decrease maximum waiting time in a program considerably, although at cost of fewer transplantations (18).

Although these papers show importance of priority rules and present theoretical suggestions for them, few investigate practices side-by-side. However, Pétér Biró and colleagues looked at the influence of the different criteria in the scores used in the UK and Spanish program (19, 20). They saw that the most influential criteria in the UK model was waiting time and in the Spanish same blood group-transplants. However, this only looked at how much weight was given by each criterion in the scores and does not study outcome measures side-to-side. As this study will focus on priorities in European KEPs a further description of these and an example demonstrating the differences between these programs will follow.

Priority rules in European kidney exchange programs

The practices of the KEPs in Europe were described in a handbook published 2019 by European Network for Collaboration on Kidney Exchange Programs (ENCKEP) (21). Table 1 is taken from the handbook and summarizes the priority rules in the European KEPs. Although there have been changes to several of the programs since published it gives an idea of the variation in the algorithm designs. Numbers in the table indicate so-called *hierarchical objectives* and their order. This means lower ranking objectives should only be considered if two or more solutions are equally ranked by the objectives higher up in the hierarchy and superior to other solutions. Some programs include parameters in a score at the bottom of the hierarchy, these parameters are indicated with *w*. To further demonstrate the variation an example will be presented below, based on the priority rules in Table 1.

	Belgium	Czech Rep. (& Austria)	Netherlands	Poland	Portugal	Spain (& Italy)	Sweden	United Kingdom
max size of solution	1	1	1	1	1	1	1	2
min lengths of the cycles	-	-	4	-	-	-	-	-
max # cycles selected	-	2	-	-	-	2	-	3
max # back-arcs	-	-	-	-	-	3	-	4
max # 2-cycles and 3-cycles with embedded 2-cycles	-	-	-	-	-	-	-	1
min# desensitisations	-	w	-	-	-	-	3	-
max HLA-matching	-	w	-	w	-	-	-	w
max DR-antigen matching in particular	-	w	-	-	-	-	-	-
min age-differences between the donors and patients	5	-	-	w	-	w	-	-
priority for paediatric patient	-	-	-	-	-	w	-	-
priority for patients not yet on dialysis	-	-	-	w	-	-	-	-
priority for highly sensitive patients	-	-	-	w	-	4	-	w
priority for O patients	-	-	-	w	-	-	-	-
priority for hard-to-match patients	3	-	3	w	w	w	2	-
priority for waiting time in KEP	-	-	-	-	-	w	-	w
priority for time on dialysis	4	-	6	-	w	w	-	-
priority for same blood-group transplants	2	-	2	-	w	w	-	-
priority for pairs with AB-donors	-	-	-	-	-	w	-	-
max # of transplant centres in (long) cycles	-	-	5	-	-	-	-	-
priority for donor-patients in the same region	-	-	-	-	-	w	-	-
min the donor-donor age differences	-	-	-	w	w	-	-	w
Constraints on length of exchanges	no	no	4	3	no	3	2	3
(Longest exchange already conducted)	3	7	4	3	3	3	n.a.	3
Constraints on length of chains	n.a.	no	4	n.a.	n.a.	no	n.a.	3
(Longest chain already conducted)	n.a.	6	4	n.a	n.a.	6	n.a.	3
providing strictly better donors for compatible pairs	n.a.	yes	yes	n.a.	n.a.	yes	yes	yes
providing strictly better donors for half-compatible pairs	n.a.	yes	n.a.	n.a.	n.a.	n.a.	yes	n.a.
altruistic chain ends in the same region where started	-	-	yes	-	-	yes	-	-

Table 1: Summary of the priority rules in European KEPs. Numbers mark hierarchical objectives and w inclusion in a score. The Swedish KEP is now part of the Scandiatransplant Kidney Paired Exchange Program (STEP). Source: ENCKEP(21)

Example of outcome variation between European programs

This example was developed together with the supervisor for this project (Tommy Andersson). Below is a pool consisting of 10 donor-recipient pairs. The arrows indicate HLA-compatibility. In Figure 3 we see the compatibility structure of the pool, containing two *disconnected subgraphs* A and B. Each of A and B contain 3 feasible cycle-exchanges, giving rise to a total of 9 possible solutions. The matching between recipient 7 and donor 8 requires AB0i-transplantation. The weights on the edges represent *matching probabilities* (MP) of the recipient, a metric used to estimate how hard a recipient is to match, used in several of the KEPs. Table 2 shows the 9 solutions together with characteristics.

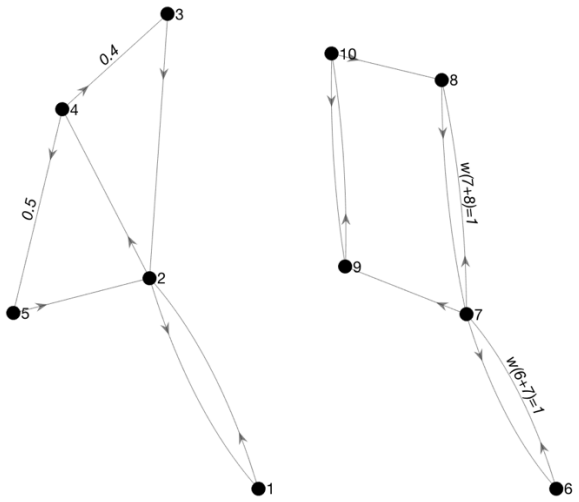


Figure 3: Compatibility graph of the pool described above. Edges mark HLA-compatibility and weights matching probability (MP). Subgraph A left, B right.

Solution	Cycles (v_i-v_j-v_i)	No. Matched Recipients	Max. cycle length (Number of cycles)	$AB0_i$ matching included
I.	2-4-3-2, 7-9-10-8-7	7	4 (2)	No
II.	2-4-3-2, 6-7-6, 9-10-9	7	3 (3)	No
III.	2-4-3-2, 7-8-7, 9-10-9	7	3 (3)	Yes
IV.	2-5-4-2, 7-9-10-8-7	7	4 (2)	No
V.	2-5-4-2, 6-7-6, 9-10-9	7	3 (3)	No
VI.	2-5-4-2, 7-8-7, 9-10-9	7	3 (3)	Yes
VII.	1-2-1, 7-9-10-8-7	6	4 (2)	No
VIII.	1-2-1, 6-7-6, 9-10-9	6	3 (2)	No
IX.	1-2-1, 7-8-7, 9-10-9	6	3 (2)	Yes

Table 2: Combinations of feasible matchings (solutions) for the example pool in Figure 4

Applying the applicable priority rules from Table 1 on the solutions with only the information above yield the possible solutions marked with X in Table 3[1]. Marked in grey is an arbitrary selection of a final solution for each program. This shows that even in a simple example all programs, except two (in this case Netherlands and Spain), can theoretically chose different solutions.

[1] Match probability is measured somewhat different between the programs but is here assumed to at least prioritize matchings the same way

Solution	Belgium	Czech Rep. & Austria	Netherlands	Poland	Portugal	Spain (& Italy)	STEP	U.K.
I.	**X**				X			
II.	X	X	X	X	X	X	**X**	
III.	X	X	**X**	X	X	**X**		
IV.					**X**			
V.		X		**X**	X			
VI.		**X**		X	X			
VII.								
VIII.								**X**
IX.								X

Table 3: Remaining solutions after applying priority rules from table 1 with the information from the example above. Cells highlighted in grey marks an arbitrary choice of solution for every program.

Purpose of the study

The example in the previous section shows the theoretical variation in how the European kidney exchange programs can choose final solution. Previous research has looked at the weight influence of the criteria in the scores of the UK and Spanish program but have been limited to these programs and has not presented differences in final solutions.

The purpose of the study was to investigate how the use of donor and recipient factors to prioritize between feasible exchanges in European kidney exchange programs impact the choice of exchanges to be carried out. The purpose was not to take stand on what priorities are better than other, but rather to investigate the how these priorities influence simulated outcomes and if this differs between programs. Through running KEP algorithms with pairwise exchanges on simulated recipient-donor pools we aimed to study this by answering the questions

- How, and to what extent, do simulated outcomes vary between different algorithms?
- How do the simulated outcomes vary with number of included pairs and increased immunization level in the underlying pool?

Methods

Theory

Program design and priority rules in kidney exchange

To understand the foundations of KEPs a brief description of program designs and algorithms follow below. Specific programs will be more thoroughly described later.

Inclusion criteria and compatibility

Traditionally only incompatible pairs (HLAi or AB0i) have been included in the programs. With desensitization protocols additional aspects of compatibility have evolved. Couples that are HLA-compatible, but AB0*i*, are referred to as *half compatible*. Whilst some KEPs only accept incompatible pairs, some include half compatible or even compatible pairs. Other programs allow *multiple* donors to be registered per patient. Some programs also include *altruistic donors.* These are donors willing to donate without connection to a recipient. This is legally forbidden in some countries like France, Poland, and Portugal (22).

Methods for finding an optimal Matching

An algorithm in a KEP first identifies all possible exchanges, then checks which of these are feasible (compatible). In pairwise exchanges this means that both recipients are compatible with the opposite donor. A set of priority rules then determine how to prioritize which of the feasible exchanges that should be carried out.

Priority rules are generally implemented as *hierarchical objectives*, a *score,* or a combination. Hierarchical objectives rule out less favorable solutions step by step in a strict order. In most KEPs the highest-ranking objective is to maximize number of transplants. Common lower ranking objectives are minimizing number of AB0i transplants, prioritize *hard-to-match* patients etc. Naturally a larger KEP will generate more solutions, thus increasing the need for lower ranking objectives to find a unique solution.

Hierarchical objectives put priorities in a strict order but do not weigh them against each other. A score puts a weight on every priority, thus demanding the designer of the algorithm to quantify difference in importance between priorities, not just to rank them. A score often increase likelihood of identifying a single final solution, although solutions in theory can have the same score. Many KEPs, for example Poland and Portugal, use a single hierarchy (maximal size solution) followed by a score.

Graphical notation

Graphical notation is typically used to demonstrate KEPs. The so-called *compatibility graph G* is a (directed) graph containing *vertices,* representing patient-donor pairs, connected with the *edges*. An edge e_{ij} exists between vertex v_i and vertex v_j, if donor d_i is compatible with recipient r_j. Each edge can be assigned a weight w based on the priority of the matching. Figure 4 exemplifies a graph using the notation above, representing a pool of 4 patient-donor pairs. Looking at the compatibility graph, we can see two solutions generating 4 transplants:

- Two *2-cycles*, matching pair 2 with 3, and pair 1 with 4
- One *4-cycle,* initiated by d_1 donating to r_2 and closing by d_4 donating to r_1.

A so-called *maximum weight matching problem* is used to find a solution. This problem formulation and notation was used for the simulation of this thesis.

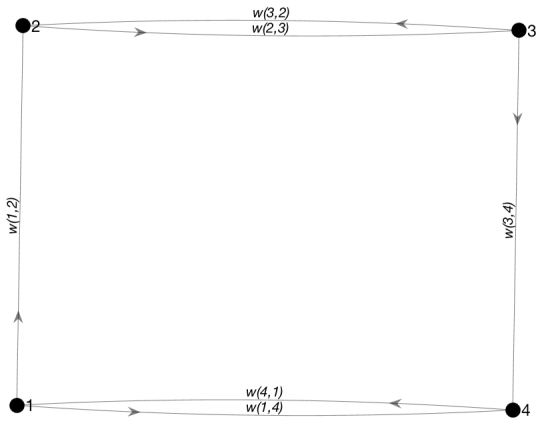


Figure 4:Example of notation in a KEP consisting of 4 patient-donor pairs.

Limitations

The simulations were limited to pairwise exchanges (no longer cycles, chains, or altruistic donation). Only one donor could be registered per recipient. No AB0i transplants were allowed. Only incompatible pairs were included in the program.

The algorithms used in the simulations were meant to represent the algorithms used in the UK, Netherlands, Spain, and Scandinavia, but modified to fit the limitations of this study and the data available.

Data selection

Donor and recipient data was collected from the Scandiatransplant database YASWA. Only data from Karolinska University Hospital donors and patients were used. All data were anonymous, and donors and patients were only identifiable by their Scandiatransplant number.

The following inclusion criteria was used:

- Living donation (LD): Recipient and donors with kidney donation/transplantation 2012-09-21 to 2022-09-21 (310 donors, 308 recipients).
- Deceased donation (DD): Recipients transplanted 2012-09-21 to 2022-09-21 (658 recipients).
- Waiting list (WL): ESRD-patients on the WL for transplantation with a deceased donor kidney 2022-09-21 (165 patients).
- STEP: Recipients and donors with kidney donation/transplantation since the start of the kidney exchange program and until 2022-09-21 (13 donors, 16 recipients).

The following clinical parameters from the database were used:

- Type of donation (DD, LD, STEP, WL)
- Age
- Sex
- AB0 blood-group
- cPRA (latest value)
- Waiting time (active) until transplantation (LD, DD, STEP)
- Waiting time to date (WL)

The aim was to generate a clinically relevant project database consisting of accepted kidney transplant patients and kidney donors. Kidney exchange is normally performed in the living kidney donation setting and the database should ideally reflect this situation. However, using only the LD-data would likely generate selection bias as this reflects recipients who has already found a matching partner. To compensate recipients transplanted with deceased kidney donors and waiting list patients were included. Patients on the waiting list and recipients transplanted with deceased donors are generally older and have longer waiting time than in the living donation setting. Also, the number of LD recipients with all available data were few (53 recipients). To compensate for this all living donation recipients were doubled in the project database.

In consultation with supervisor and transplant surgeon Lars Wennberg, 13 patients (10 WL and 3 LD) with cPRA<2 and active waiting time>1500 days were removed as this data was assumed to be wrong inputs.

After adding together all data, removing duplicates and then duplicate all LD-recipients, all recipient data without combined cPRA and active waiting time was removed. 964 rows were removed and 462 remained (106 LD, 207 DD, 149 WL). The reason for the large shortfall was mainly that combined cPRA was introduced recently and therefore most deleted datapoints were older than 4 years. Hence, risk for bias should be limited, despite shortfall.

In the project database the average donor age was 48.6 years and recipient age 51.6 years, average waiting time among patients was 322 days and average cPRA 20.2. The share of blood group 0 individuals was 40.0% among patients and 39.4% among donors. 36.8% of patients were female, whilst 61.0% among donors. Histograms and scatter plots for distribution of the main parameters in the project database can be found in the Appendix 1.

Parameter modification and assumptions

Waiting time

Waiting time, as used in KEP algorithms, refer to the time enlisted in a program or on a waiting list at the time for a matching round. For DD and LD recipients in the database waiting time is the time until they got their transplantation. To compensate for this, waiting time was redistributed for these patients to better reflect a feasible set of waiting times at the time of a matching round. The redistributed waiting time $waitingtime^*_i$ was generated using the original waiting time $waitingtime_i$ according to

$$waitingtime^*_i = waitingtime_i * 0.1 * c$$

Where c is a random integer belonging to [1,10]. This gives a uniform redistribution of the waiting time, with expected value

$$E(waitingtime^*_i) = 0.5 * waitingtime_i$$

Spouse donors and cPRA

Women with spouse donors have higher probability of a positive crossmatch with that specific donor than reflected in *cPRA*. Approximately 28% of transplanted women (in Sweden) have a spouse donor (23). For women assumed to have spouse donors, we

used the same assumption as in previous literature by Alvin Roth(24), such that the probability for a crossmatch with a spouse donor cPRA^*_i was:

$$cPRA_i^* = 25 + cPRA_i * 0.75$$

HLA-mismatch

HLA typing was not available for a large amount of data. As HLA-mismatch level is relatively random at nature between two non-related individuals (such as when being matched in a kidney exchange) we assumed the HLA-mismatch level between any patient and donor in the pool to be drawn at random from the distribution of HLA-mismatches in transplantation in UK (25). The reported mismatch level depends on difference in HLA- A, B and DR, for exact definition please see reference[2]. *Level 1* corresponds lowest level of mismatch.

$$\begin{cases} P(Level1) = 0.03 \\ P(Level2) = 0.29 \\ P(Level3) = 0.43 \\ P(Level4) = 0.25 \end{cases}$$

To present the example some notation will be needed.

Simulations

The simulations were carried out in MATLAB R2022a. The basic simulation structure was taken from a similar simulation study by Tommy Andersson and Jörgen Kratz (14), downloaded at the online publication site[3]. This was then modified to fit the purpose of this study. The Scandinavian algorithm was already coded for, but the other algorithms were implemented in a similar manner to the Scandinavian code.

The following 4 simulations were carried out

- *Base case*: Number of patient/donor pairs drawn from the project database (n) equal to 150, number of iterations/matching rounds simulated (N) equal to 50
- *Increased number of pairs*: n=250, N=50
- *Decreased number of pairs*: n=50 N=50
- *Increased cPRA in database:* Same as base case, but doubled all patients with cPRA over 20% in the project database (134 patients)

[2] https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/23470/section-5-kidney-activity.pdf
[3] https://doi.org/10.1093/restud/rdz018

Simulation structure

The simulation was carried out based on 4 inputs:

- Number of simulated patient-donor pairs *n*
- Number of iterations *N*
- .xls-file containing the above specified patient data
- .xls file containing the above specified donor data

n donors and *n* patients were randomly selected from the project database (with replacements). These were paired together, giving *n* initial patient-donor pairs. In this initial matching AB0 compatibility was compared withing the pairs, and a *simulated crossmatch* was carried out. The simulated crossmatch used c*PRA** / cPRA as the probability of a positive crossmatch. The crossmatch result was drawn randomly from this probability distribution. As mentioned, c*PRA** represents the probability of a positive crossmatch between a female patient and her spouse donor. 28% of transplanted female recipients in Scandinavia have spouse donors. With 39% male (or unknown sex) donors in our project database, 70% of male donors matched with a female in the initial matching were therefor assumed to be spouses (23).

After the *simulated crossmatch* and test for AB0-compatability within the initial pairs, compatible pairs went on to be transplanted i.e., *self-matched* and incompatible pairs were included in the KEP-matching round.

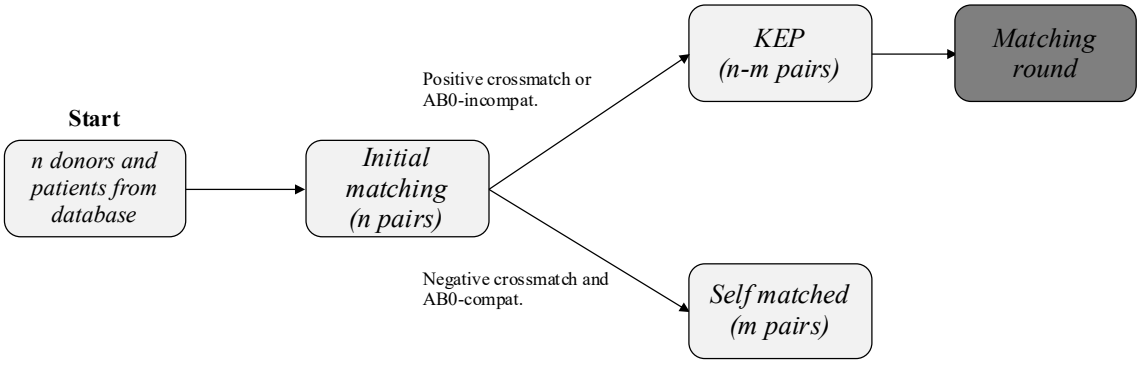


Figure 5: Schematic overview of first simulation steps: Randomized selection of pairs, initial matching, and exclusion for self-matching

The matching rounds (the KEP)

At first the compatibility between the pairs included in the KEP was tested. This was done with simulated crossmatching (as described earlier) and test for AB0-compatability between all the donors and patients from the included pairs. All feasible (compatible) exchanges were represented by an edge to build a compatibility graph.

Weights were assigned to the edges according to the priority rules in the studied programs. For a schematic view, please see Figure 6. The algorithms and how weights were assigned are to be specified later.

The maximum weight matching problem was solved using an implementation of the so-called *Blossom algorithm* and the *primal-dual for maximum weight matching problem*, published by Zvi Galil in 1986 and ported to MATLAB from Python by Daniel Saunders.(26, 27).

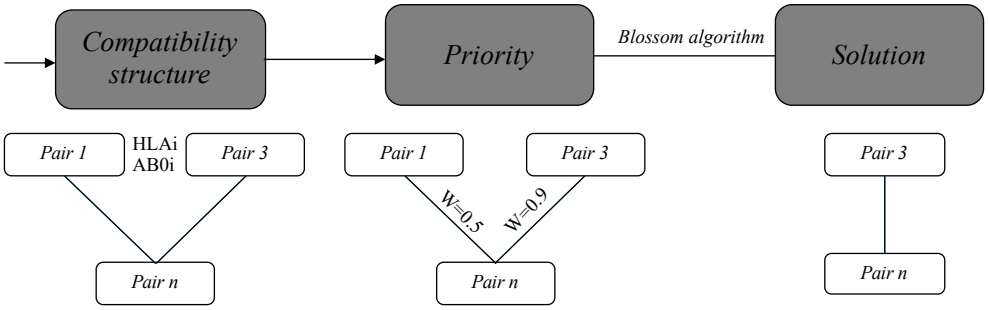


Figure 6: Schematic overview of the simulation steps in the matching round

Assigning the weights

For the solving algorithm to function (due to the maximum weight matching formulation) both values for the hierarchical objectives and scores had to be assigned as an edge weight. This is easy for a score, but also possible hierarchies. The weight contribution from different hierarchical objectives were set so that the weight added from lower ranking objectives should only matter if the solutions were equal according to higher ranking objectives. The sum of the edge weight contributions from a lower ranking hierarchical objective were assigned to be less than the minimum element in the weight contribution from a higher-ranking objective. With W_2 being a matrix with the weight contributions from a lower ranked objective and W_1 from a higher-ranking objective, we can formulate this as

$$\sum W_2 < \min W_1$$

The max weight problem can then be formulated as below. w_{1ij} is the weight contribution from objective 1 for matching recipient i with donor j, x_{ij} denotes if recipient i is matched with donor j.

$$\max \sum_{ij} (w_{1ij} + w_{2ij} \dots) x_{ij} = \max \sum_{ij} w_{ij} x_{ij}$$

$$x_{ij} \in [0, 1]$$

The solving function used is designed for undirected graph, but since the simulation only considered pairwise exchanges this is equivalent to summing the weights on the directed edges between two nodes if, and only if, they are both non-zero weights.

Implementation of the KEP algorithms

Below follows a description of the exchange programs in the UK, Netherlands, Spain, and Scandinavia and how they were used in our simulations. For full description including technical details see Appendix 2. The Dutch algorithm was retrieved from the ENCKEP handbook 2019 (21), the British and Scandinavian from the websites of the organizations responsible for the respective program (28, 29), and the Spanish via direct communication with experts at the responsible organization (30).

As previously mentioned only 2-way exchanges and no transplantations over the blood group barrier were considered. Also, only incompatible pairs were included in the program. Therefore, no priority rules concerning longer cycles, chains, AB0i transplants or compatible pairs were used. Following this, the highest-ranking objective for all programs was to maximize the number of transplants.

Scandinavia- STEP

Second to maximizing the number of transplants, the Scandinavian algorithm prioritize patients with low *Match Probability (MP)*. This measures how hard a patient is to match, depending on the patients cPRA and the compatibility structure of the pool. Few compatible donors and high cPRA yield a low MP. In the Scandinavian program this is defined as

$$MP_i = (1 - cPRA_i) * \frac{No.\,compatible\ donors\ in\ pool}{No.\,total\ donors\ in\ pool}$$

Netherlands

The Dutch KEP also uses a hierarchy, but with more levels than the Scandinavia. After maximizing the number of transplants, priorities come in the following order:

- Maximize the number of blood type identical transplants
- Match the patients in order based on *Match Probability* (lowest first)
- Match the patient with the longest waiting time

Unlike Scandinavia, the second hierarchy is maximizing the number of blood group identical transplants. The reason for this is to ensure a fair allocation for disadvantaged blood groups, such as blood group 0 (21). Match probability is here defined slightly differently, considering the compatible share of AB0-compatible donors rather than the total number of donors

$$MP_i = (1 - PRA_i) * \frac{No.\,compatible\ donors}{No.\ AB0 compatible\ donors}$$

The Dutch program also considers the spread over transplant centers, which is not included here.

United Kingdom

Unlike the previous two, the UK program uses a score second to maximizing the number of transplants. The score involves 4 parameters

- Waiting time (50p per number of previous matching rounds without a match)
- Sensitization points (cPRA divided by 2)
- HLA- mismatch (0-15 points depending on *HLA-mismatch level*)
- Donor to donor age difference (3 p if under 20 years)

HLA mismatch refers to the level of mismatch in HLA (loci *HLA-A*, *-B and -DR)* between the donor and recipient. This, and the use of donor-to-donor age difference, is unique to the UK program. We also see that the priority for sensitized patients is slightly different from the Scandinavian and Dutch program.

Spain

Like the UK program, Spain uses a score second to maximizing the number of transplants. The score involves following parameters

- Waiting time (30p if over 1 year in the program)
- Match Probability (0-30 p)
- Same blood group transplants (30p if same, 30p additional if blood group 0)
- Pediatric recipient (500p if under 18 years of age)
- Age difference between recipient and donor (0-15p)

The last parameter only gives points if the donor is older than the recipients and it is subject to a number of additional conditions as well (see Appendix 2). Match Probability is defined the same way here as in the Dutch program. Both the priority of pediatric recipients and the age difference between recipient and donor is unique to the

Spanish program. In the score there is also a parameter for time on dialysis. However, as this data was not available to us it was not used in this study.

Analysis of output

These following output parameters were analyzed for every iteration (matching round):

- Number of self-matched pairs and number of pairs included in the KEP/matching round
- Share of matched recipients that were the same across all programs
- Share of blood group 0 patients that was matched
- Share of pediatric patients that was matched
- Mean age of matched recipients
- Mean age difference between matched donors and recipients
- Mean age difference between donors in matched pairs
- Mean waiting time (redistributed) of matched recipients
- Mean cPRA of matched recipients
- Mean HLA-mismatch level for matched pairs

For most parameters the samples were the parameter means for every matching round, hence the sample size was 50 (N=50). Using the central limit theorem, these samples could be assumed approximately normally distributed. For these parameters 95% confidence intervals were reported. However, the sample shares (pediatric and blood group 0 patients being matched etc.) could not be assumed normally distributed. Hence, only means were reported (histograms Appendix 1).

To test for significant differences comparison of confidence intervals and two-sample two-tailed t-tests were used. The null hypothesis tested for was that the samples come from distributions with the same mean. The interpretation of the confidence interval comparison is that non overlapping confidence intervals imply we can reject the null hypothesis, at significance level 0.05. However, this is not equivalent. If tested with a two-sample t-test, significance may be present even with overlapping confidence intervals (type 2 error). T-tests can however only be done for the parameters assumed approximately normally distributed. As carrying out t-tests for all simulations would generate 144 pairwise comparisons, this was only done for the base case. Hence comparison of confidence intervals was used for the other simulations, with risk for type 2 errors.

Ethics

The ethical considerations in this thesis are mainly due to the use of patient data. No experiments on patients were carried out and no direct patient contact was mas made. The patient data was taken from the Scandiatransplant registers. Only data from Karolinska University Hospital, where supervisor Lars Wennberg is the Medical Director, was used. No data handled by the author could risk identifying patients or donors. Patients and Donors were only identifiable via their Scandiatransplant number.

Although ethical considerations are important in KEPs, the purpose of the thesis was not to take stand on what priorities are better than other, but rather to investigate the how these priorities influence simulated outcomes and if this differs between programs.

Results

Below the results are presented, for discussion regarding in-data and exclusion we refer to the data-section in the methods chapter.

Matching ratios and overlap

The average matching ratio in the KEPs was 38.5% in the base case. Since all programs here have maximum number of transplants as first priority, matching ratio was the same for all programs. On average 72.6% of the matched recipients were the same across programs.

Table 4Matching data for the simulations is seen in Table 4. *Self-matched* refers to the percentage of pairs that are matched in the *initial matching* with their original donor, i.e. not included in the KEP. Matched in KEP refers to the matching ratio for the pairs included in the KEP. Total matched is the percentage of all patients in the simulation (*n)* that are chosen for a transplantation i.e., either self-matched or matched in KEP. Matching overlap is the share of matched recipients that were the same across programs. Matching overlap was relatively stable in our 4 cases, although cases with lower *n* had higher variance.

Simulation	Self-matched (% *of n*)	Included in KEP (% *of n*)	Matched in KEP (% *of included*)	Total matched (% of n)	Matching Overlap (%)
n=150 (Base case)	49.0%	51.0%	38.5%	68.6%	72.6%
n=250	49.7%	50.3%	43.5%	71.6%	73.1%
n=50	49.8%	50.2%	26.3%	63.2%	74.4%
n=150 (Incr. PRA)	42.7%	57.3%	42.9%	67.3%	74.1%

Table 4: Matching data (sample averages) for all four simulations. Matching overlap refers to percentage of matched recipients that was the same across all programs.

The base case

T-tests in the base case showed significant differences for 5 of the 6 tested parameters. UK showed highest average waiting time and lowest HLA mismatch levels.

Scandinavia (STEP) showed highest average cPRA and Spain lowest average age among matched recipients. Table 5 summarizes the significant differences from the t-tests (dotted lines) and Figure 7 show plots with means and confidence intervals.

Waiting time*	**cPRA**	**HLA-mismatch**	**Age (recip.)**	**Age difference (donor-recip.)**
UK *(highest)*	Scandinavia *(highest)*	UK *(lowest)*	Spain *(lowest)*	Spain *(smallest)*
Spain	Netherlands	Spain	Scandinavia	Scandinavia
Netherlands	UK	STEP	UK	UK
Scandinavia	Spain	Scandinavia	Netherlands	Netherlands

*Table 5: Order om the sample means. Dotted lines mark significant differences (t-test at significance level 0.05). Note that scales go in different directions. *redistributed waiting time*

The difference between highest average waiting time among matched recipients (UK) and lowest (Scandinavia) was 54 days. T-test showed significant difference between all programs except Netherlands and Scandinavia. The difference with lowest significance was between UK and Spain ($p = 0.043$).

The difference in average cPRA between highest (Scandinavia) and lowest (Spain) was 8.2 percentage points. T-tests showed significant difference between all programs except UK and Netherlands. The order between the programs was different for waiting time and cPRA which at first might seem contra intuitive, as there is likely correlation between cPRA and waiting time for kidney transplantation (appendix 1). However, when plotting cPRA and waiting time for the recipients in unique matchings (matching that is unique to one program) there are high-cPRA recipients with shorter waiting time (<300) that seem to be prioritized in the Dutch and Scandinavian program. At the same time there are several recipients with low c-PRA and longer waiting time (>300) that tend to get transplanted in the UK and Spanish Programs.

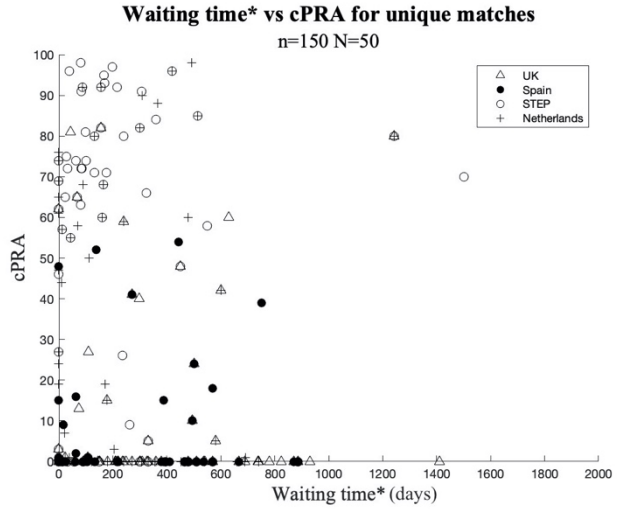


Figure 7: Waiting time vs cPRA for unique matches. To the right blood group distribution (share of total number of unique matches) for unique matches. (waiting time redistributed according to assumption)*

Difference in HLA mismatch level 0.16 (scale 1-4, definition in appendix 2) between lowest (UK) and highest (Netherlands). T-tests showed significant difference between UK and all other programs.

The difference in average age between the highest (Netherlands) and lowest average (Spain) was 2.1 years. T-tests and confidence intervals showed lower average age among matched recipients for Spain than the other programs. Average age difference between donors and recipients was close to 0 in the Spanish program, whereas donors in other programs were on average younger. Difference between largest average age difference (Netherlands) and Spain was 2.2 years. Matching ratio for pediatric patients was higher in Spain than remaining countries, 33 percentage points higher than program with lowest matching ratio (Netherlands).

The average matching rate of blood group 0 patients was 18-20%, but no substantial difference could be seen across programs. No difference could be seen in age difference between donors in matched pairs either.

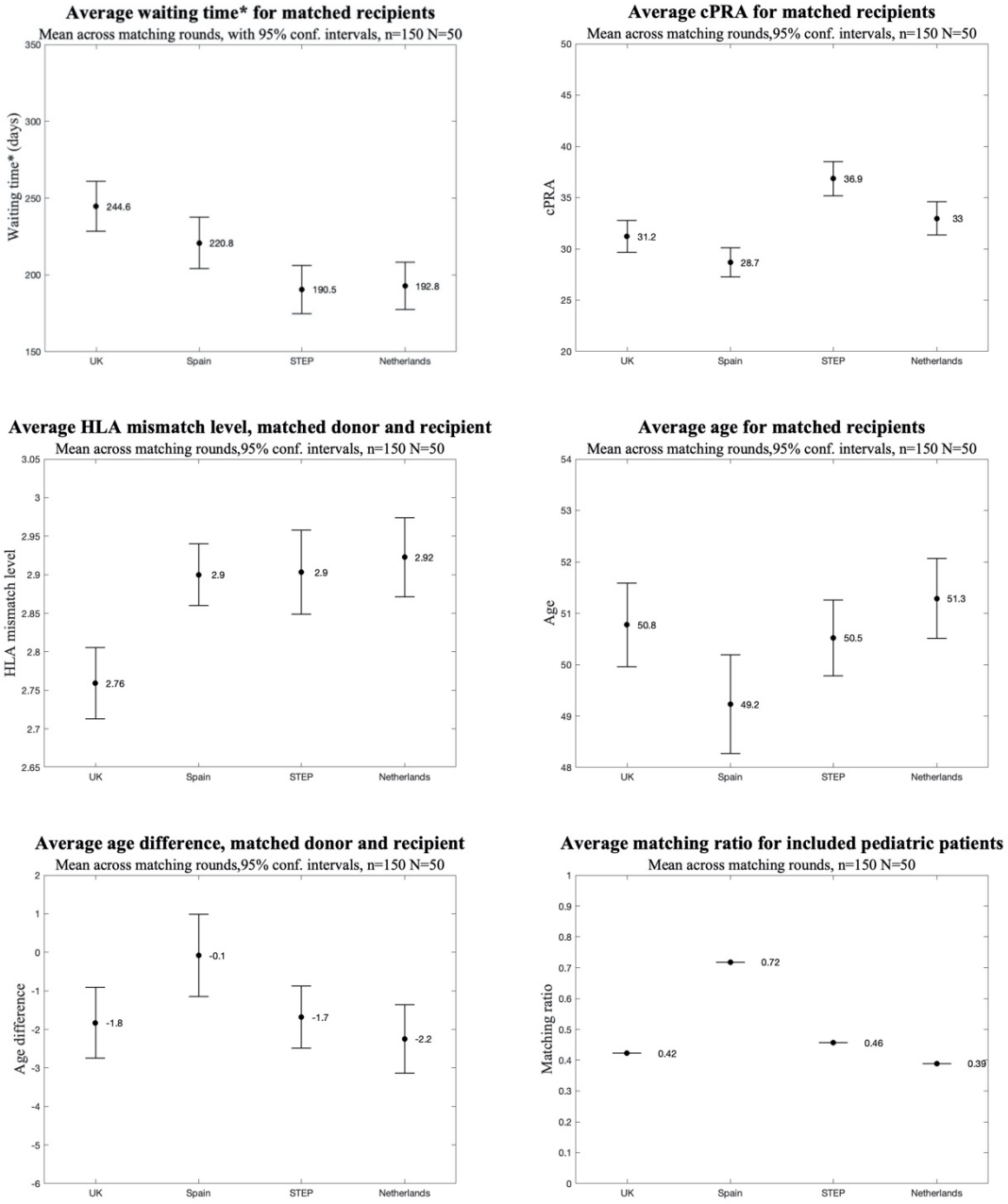


Figure 8: Averages and 95% confidence intervals (error bars) for the base case. UK left, Spain middle left, Scandinavia (STEP)) middle right, Netherlands right in each.

Increased number of pairs (n=250)

Most confidence intervals narrowed when increasing the number of pairs, confirming trends seen in the base case. The numeric differences and difference between highest and lowest average was relatively constant, except slightly larger difference in HLA-mismatch and smaller in average age. All plots, and additional tables are found in appendix 1

In Figure 9, we see the trend of narrowed confidence intervals for cPRA. Differences in average cPRA were accentuated, with only UK/Netherland confidence intervals overlapping. Order between programs was the same, with highest in Scandinavia. In terms of waiting time the difference between UK and Spain decreased, whilst difference down to Scandinavia/Netherlands increased. HLA-mismatch difference was slightly increased, emphasizing UK as having lowest average HLA-mismatch. Average matching ratio for pediatric patients increased, but difference between programs was constant. Confidence intervals for the average age among recipients narrowed, but difference between programs remained approximately same. Slight increase in age difference between donors was seen, with UK having lowest mean, however confidence intervals still had overlap.

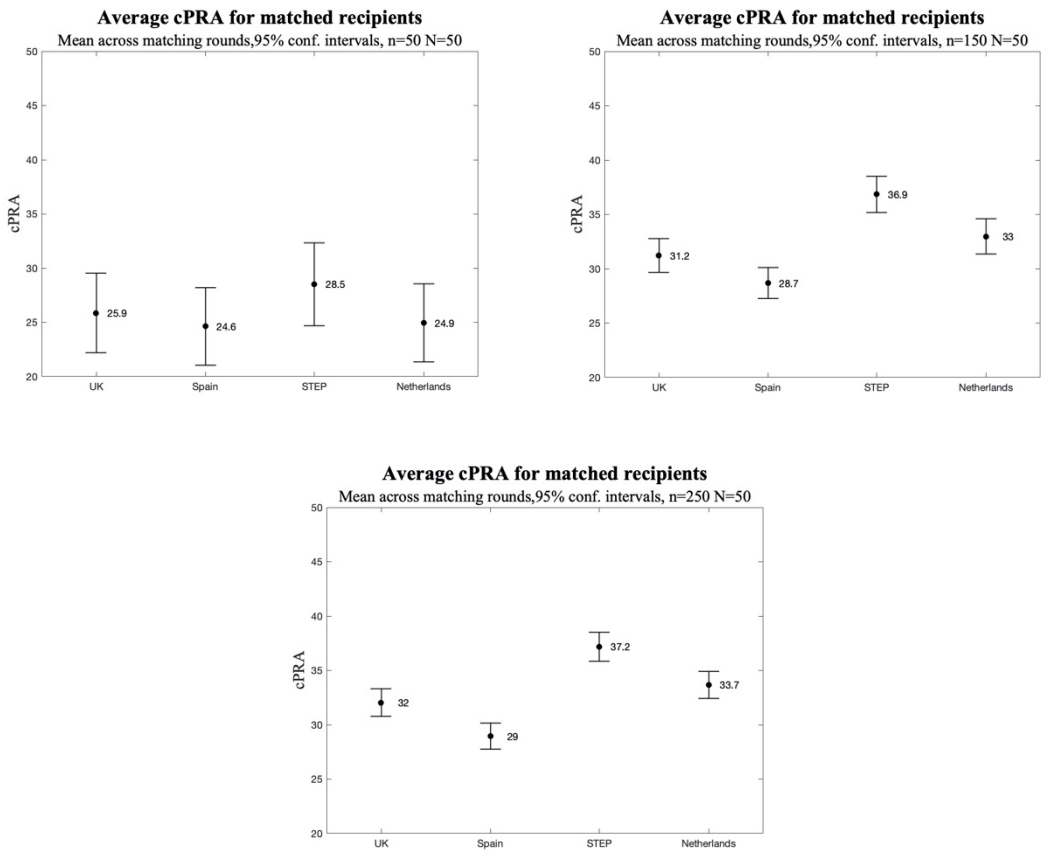


Figure 9: Average cPRA with 95% confidence intervals for decreased case (n=50) top left, the base case (n=150) top right and increased case (n=250) bottom.

Decreased number of pairs (n=50)

When decreasing the number of pairs there was overlap in all confidence intervals and differences between programs decreased (Table 6 Appendix). The difference between highest and lowest mean was decreased for all parameters. Although not tested statistically there was a noticeable shift towards lower average cPRA (see Figure 9), matching ratio for blood group 0 and pediatric patients.

Although confidence intervals were overlapping for all parameters, some of the same tendencies as in the base case could be seen. Scandinavia had slightly higher sample cPRA, while Spain and UK had slightly higher sample waiting time. This could be seen when looking at unique matches as well. UK had slightly lower average HLA-mismatch than others. All plots for this simulation, similar to Figure 7, can be found in appendix 1.

Increased cPRA in project database (n=150)

In the increased cPRA case there was higher average cPRA among the matched recipients across programs (Figure 10). Confidence intervals for some parameters were narrower, probably reflecting the higher number of matches (inclusion and matching data in Table 4). For age among recipients and waiting time, there was slightly less difference between the programs. Few other differences could be seen for the increased cPRA case compared to the base case.

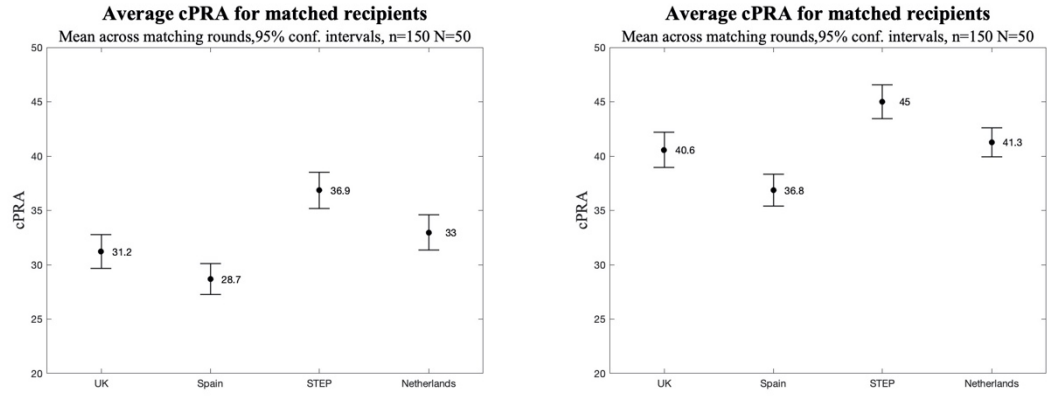


Figure 10: Average cPRA for matched recipients in the increased cPRA case (to the left), and the base case (to the right). UK left, Spain m.left, Scandinavia (STEP) m.right, Netherlands right

Discussion

Our findings support that there are systematic differences in how the studied programs match pairs in pairwise kidney exchanges, at least when the number of included pairs is large enough. 72.6% of matched recipients in our base case were the same across programs. This number was approximately constant when changing number of included pairs. With a small pool size little systematic difference in outcome was detectable, but when increasing pool size the differences became clearer and, in some cases, increased. In our base case we could see significant differences in 5 of the 6 parameters tested for. When increasing the sensitization level in the pool, little additional difference was seen except increased sensitization among matched recipients. As expected, matching ratio in the KEPs went up as the number of included patients increased. It was somewhat surprising that matching ratio (in the KEPs) was higher in the increased PRA-case than the base case. This hints that the effect of higher number of included patients was stronger than the negative effect of higher average cPRA in this case.

While all programs include some priority rule for sensitized patients, our base case suggests the Scandinavian program match recipients with the highest cPRA. As this is the only priority except maximizing number of transplants in the Scandinavian program, whilst other programs have this priority lower down in the hierarchy, this is understandable. Spain, and especially UK, had higher average waiting time among recipients than remaining programs. This is also reasonable as Spain and the UK include priority for waiting time at 2nd level in the hierarchy (when only considering pairwise exchanges), the Netherlands at the 4th level and Scandinavia does not have a priority for waiting time. For HLA mismatch, UK is the only program with priority rules for HLA mismatch level, which was also seen in the results. However, the absolute difference was relatively small (0.16) and the scale (level 1-4) is somewhat hard to interpretate. The Spanish program showed lower average age among recipients. The Spanish program also had smaller age difference between donors and recipients. Since donors are generally younger than recipients this might contribute to the lower average age. Looking at the matching ratio of pediatric patients Spain had higher ratio than the other programs. This is understandable as the program is the only one with priority rules for pediatric patients, and might also contribute to the lower average age.

However not all priority rules had substantial effect on outcome. The Spanish and Dutch programs have priority rules for identical blood group transplants, and the Spanish has additional rules for prioritizing blood group 0 patients. These rules are

likely implemented to increase matching ratios for blood group 0 patients. Even though the Spanish and Dutch program had higher average matching ratio for blood group 0 patients in all cases, the effect was small (1-3pp). The UK algorithm has priority rule for limiting age difference between donors in matched pairs and although this is unique to the UK program no significant effect was seen.

Comparing our results to previous studies the matching ratios seen in our simulations are close to, but slightly lower, than in Andersson and Kratz (14). This is natural as the basic simulation structure is shared, but input data is different. Another similar simulation study with pairwise exchanges, but dynamic exchange pools (hence varying pool size), found a matching ratio of 41% (31). This compares to 38.6% in our base case. These similarities support the setup of our simulation but does not compare our results regarding the output or differences across programs. In Biró et al (21), the criterion contributing to most of the weight in the UK score was waiting time. The least influential criterion was donor-donor age difference. This harmonizes with our results. On the other hand, their results for the Spanish program partly contrast our findings. They found blood group priority to contribute with most weight and match probability second. However, there are some possible explanations for this. Firstly, input data in their study was different (did not involve pediatric patients) and pool sizes and specific details on the algorithm might be different as they were not revealed. Secondly, weight contribution does not necessarily give effect on outcome. It only does so if matchings are conflicting and is therefore depending on compatibility structure in the pool. Similarly, matching ratio for blood group 0 patients in our study was not substantially higher in the Dutch program either, although one could expect influence to be high as priority for same blood group transplants is high up in their hierarchy.

Few other comparable studies exist to our knowledge. A systematic overview from 2021, studying reporting standards in European kidney exchange programs, found 5 articles published after 2015 studying outcomes in European KEPs (32). These articles, annual reports from the organizations responsible for the national KEPs, and additional papers from searches in various databases were read but limited comparable studies and no side-to -side comparisons of outcomes in European KEPs could be found. The limited data was concentrated to the Netherlands. In a report, matching ratio for blood group 0 patients within the Dutch program was 20% for a pool size of 78 in 2021 (33). Similarly, a simulation study with pairwise exchanges in the Dutch program showed a total matching ratio of 47% and 21% for blood group 0 patients, with a pool size of 312

(34). Even if these figures support some of our findings, this is a small share and does not speak about the difference between programs. With limited comparable research and limited consensus on some areas, we can conclude that further research is needed on the topic.

The main strength in of this paper is that it is a side-to-side comparison using mainly real word donor and recipient data. Real world- data allows for capturing realistic distributions and correlations. The side-to-side comparison lets us compare outcome parameters in an efficient manner as all is based on the same input. Also, the simulation setting captures some of the natural randomness, whilst allowing for high number of iterations thus easily capturing systematic differences.

However, there are four main areas to consider when interpreting the results. Firstly, simulations are limited to pairwise exchanges within the blood group barrier. Allowing longer cycles (or chains) would likely increase number of feasible matchings, but not the parameter distributions in the pool. This is partly similar to increasing pool size as covered in this paper. However, allowing for AB0i in the way most programs use it would likely change blood group distribution and increase average cPRA in the pool, thus possibly influencing the results. This is partly covered in our increased cPRA case, but still limits generalization of results. Secondly the algorithms are not exact replicas. Some priority rules (like time on dialysis in the Spanish program) are not implemented in our simulations due to the lack of data or project limitations. How this would affect results is hard to tell but as time on dialysis is likely connected to long waiting time, this may further prioritize these patients in the Spanish program. Thirdly it should be noted that the project database does not contain donors and patients directly subject to kidney exchange, but a mixture of waiting list patients and previously transplanted recipients and donors combined to imitate the KEP setting. Due to practical reasons the data was only taken from one transplant center (Karolinska University Hospital). As distributions might depend on geography and KEPs are tailored to their own setting, it should be questioned if this can be cause for bias. However, the scoop of this thesis was to compare allocation priorities side to side based on the same input and not generalizing to KEPs actual performance in their natural setting.

Conclusions

The main conclusion was that our findings support that differences in prioritization rules between the studied KEPs also impact which exchanges are carried out, at least in the setting of pairwise exchanges within the blood group barrier. The most prominent differences in our study were that the British KEP matched recipients with highest average waiting time, the Scandinavian with the highest average cPRA and the Spanish matched the highest ratio of pediatric patients. Differences across programs was clearer with increasing number of included patients. These differences did not change substantially with increasing cPRA in the underlying population. It should be noted that this study has some limitations influencing ability to generalize results, and therefore call for additional research.

In the short term these findings serve as motivation for further research to investigate these differences, as little research is done comparing the current practices. For the longer term our results imply that differences between programs become more systematic as programs grow. As KEPs are likely to grow through increased pool sizes and international collaborations, the importance of carefully considering prioritization rules and discussing between programs becomes clear.

Author's Contribution

The design of the study was developed together with the supervisors. The simulation code was based on the code used in a paper written by supervisor Tommy Andersson and colleague Jörgen Kratz (14). The Spanish, British and Dutch algorithms were implemented by the author with some support from Andersson and Kratz and the structure was tailored to this specific project. The input data was chosen and discussed together with supervisor Lars Wennberg. All analysis and output were generated by the author, but discussed with supervisors.

References

Uncategorized References

1. Yang F, Liao M, Wang P, Yang Z, Liu Y. The Cost-Effectiveness of Kidney Replacement Therapy Modalities: A Systematic Review of Full Economic Evaluations. Applied Health Economics and Health Policy. 2021;19(2):163-80.
2. Wolfe RA, Ashby VB, Milford EL, Ojo AO, Ettenger RE, Agodoa LYC, et al. Comparison of Mortality in All Patients on Dialysis, Patients on Dialysis Awaiting Transplantation, and Recipients of a First Cadaveric Transplant. New England Journal of Medicine. 1999;341(23):1725-30.
3. Eurotransplant. 2052P_All ET_kidney. 2021.
4. Scandiatransplant. Historical Data: Transplantation Figures. 2021.
5. Scandiatransplant. Historical Data: Waiting list figures 2021.
6. Tydén Gea. Blodgruppsinkompatib-la njurar kan transplanteras. Läkartidningen. 2012(39 2012).
7. Joher N, Matignon M, Grimbert P. HLA Desensitization in Solid Organ Transplantation: Anti-CD38 to Across the Immunological Barriers. (1664-3224 (Electronic)).
8. Althaf MM, El Kossi M, Jin JK, Sharma A, Halawa AM. Human leukocyte antigen typing and crossmatch: A comprehensive review. (2220-3230 (Print)).
9. Biró P, Haase-Kromwijk B, Andersson T, Ásgeirsson EI, Baltesová T, Boletis I, et al. Building Kidney Exchange Programmes in Europe-An Overview of Exchange Practice and Activities. Transplantation. 2019;103(7):1514-22.
10. NHS NHSU. Organ donation and transplantation activity report 2018-2019. Transplantation Ba; 2019.
11. ONT. Actividad de donación y transplante renal espana 2019. Organizacion Nacional de Transplantes (ONT); 2019.
12. NTS. NTS cijferoverzicht. Nederlandse Transplantatie Veriniging (Dutch Transplant Foundation); 2019.
13. Scandiatransplant. Transplant and waiting figures. 2019-2021.
14. Andersson T, Kratz J. Pairwise Kidney Exchange over the Blood Group Barrier. The Review of Economic Studies. 2019;87(3):1091-133.
15. Øien CM, Reisæter AV, Leivestad T, Dekker FW, Line PD, Os I. Living Donor Kidney Transplantation: The Effects of Donor Age and Gender on Short- and Long-Term Outcomes. Transplantation. 2007;83(5).
16. Stepkowski SM, Mierzejewska B, Fumo D, Bekbolsynov D, Khuder S, Baum CE, et al. The 6-year clinical outcomes for patients registered in a multiregional United States Kidney Paired Donation program - a retrospective study. Transplant International. 2019;32(8):839-53.
17. Liu W, Melcher ML. Matching Kidneys with Priority in Kidney Exchange Programs. Medical Decision Making. 2019;39(4):336-46.
18. Monteiro T, Klimentova X, Pedroso JP, Viana A. A comparison of matching algorithms for kidney exchange programs addressing waiting time. Central European Journal of Operations Research. 2021;29(2):539-52.
19. Druzsin KaB, Péter and Fleiner, Rita and Klimentova, Xenia. Simulations for measuring efficiency of international kidney exchange programmes. 16th International Symposium on Operational Research in Slovenia. 2021.
20. Klimentova KoDPBRFX. Large scale performance analysis of international kidney exchange programmes by the ENCKEP simulator. 9th VOCAL Optimization Conference:

Advanced Algorithms. 2022.
21. Andersson T, Biró P, Calderón M. Modelling and Optimisation in European Kidney Exchange Programmes - Second Hnadbook of Working Group 1 in the COST action CA15210: European Network for Collaboration on Kidney Exchange Programmes (ENCKEP). ENCEKP; 2019.
22. Biró P, van de Klundert J, Manlove D, Pettersson W, Andersson T, Burnapp L, et al. Modelling and optimisation in European Kidney Exchange Programmes. European Journal of Operational Research. 2021;291(2):447-56.
23. Ingela Fehrman-Ekholm, Annette Lennerling, Niclas Kvarnström, Magnus Rizell, Olausson M. Transplantation av njure från levande givare – en framgångssaga. Läkartidningen. 2011.
24. Roth AE, Sönmez T, Ünver MU. Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences. American Economic Review. 2007;97(3):828-51.
25. UK NHSN. Kidney Activity Report. National Health Services (NHS), Transplant Ba; 2021.
26. Saunders D. Weighted maximum matching in general graphs MATLAB Central File Exchange.2022.
27. Galil Z. "Efficient Algorithms for Finding Maximum Matching in Graphs". ACM Computing Surveys. 1986.
28. NHS NHSU. POL274/5 – Living Donor Kidney Transplantation. 2022.
29. STEP-programme S-candiaT-ransplant kidney E-xchange P-rogramme (STEP), Version 1.11. Scandiatransplant; 2019.
30. Torres EC. SNKEP info. In: Lundgren S, editor. E-mail correspondance regarding SNKEP algorithm ed2022.
31. Santos N, Tubertini P, Viana A, Pedroso JP. Kidney exchange simulation and optimization. Journal of the Operational Research Society. 2017;68(12):1521-32.
32. Smeulders B, Mankowski MA, van de Klundert J. Kidney Exchange Program Reporting Standards: Evidence-Based Consensus From Europe. 2021(2296-2565 (Electronic)).
33. NTS. NTS cijferoverzicht. Nederlandse Transplantatie Veriniging (Dutch Transplant Foundation); 2020.
34. De Klerk M, Van Der Deijl WM, Witvliet MD, Haase-Kromwijk BJJM, Claas FHJ, Weimar W. The optimal chain length for kidney paired exchanges: an analysis of the Dutch program. Transplant International. 2010;23(11):1120-5.
35. Keizer KM, de Klerk M, Haase-Kromwijk BJJM, Weimar W. The Dutch Algorithm for Allocation in Living Donor Kidney Exchange. Transplantation Proceedings. 2005;37(2):589-91.
36. Coll E. E-mail correspondance. . Discussion about updates since the ENCEP 2019 handbook ed2022.

Appendix 1

Results, additional data

Below are additional plots and tables for all simulations.

Difference between highest and lowest parameter means

Parameter	***n=50***	***n=150***	***n=250***
Waiting time (days)	*34.2*	*54.1*	*54.4*
cPRA (%)	*4.1*	*8.2*	*8.2*
Age (yrs.)	*1.2*	*2.1*	*1.6*
HLA mismatch (level)	*0.09*	*0.16*	*0.21*
Age Difference D-R (yrs.)	*1.3*	*2.1*	*2.0*
Age Difference D-D (yrs.)	*1.0*	*0.7*	*0.9*
Pediatric patients (%)	*13*	*33*	*33*
Blood group 0 patients (%)	*2.0*	*2.0*	*2.0*

Table 6: Difference between the program with highest and lowest parameter mean for every simulation

Input data analysis (from project database)

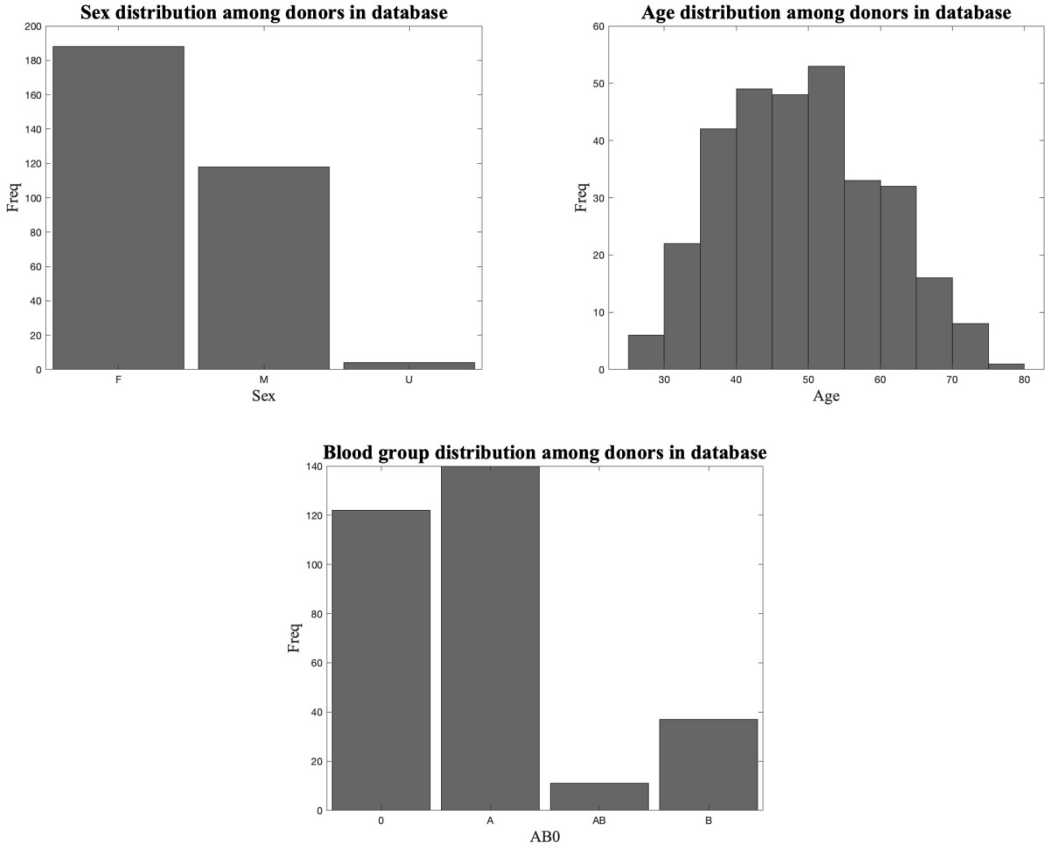
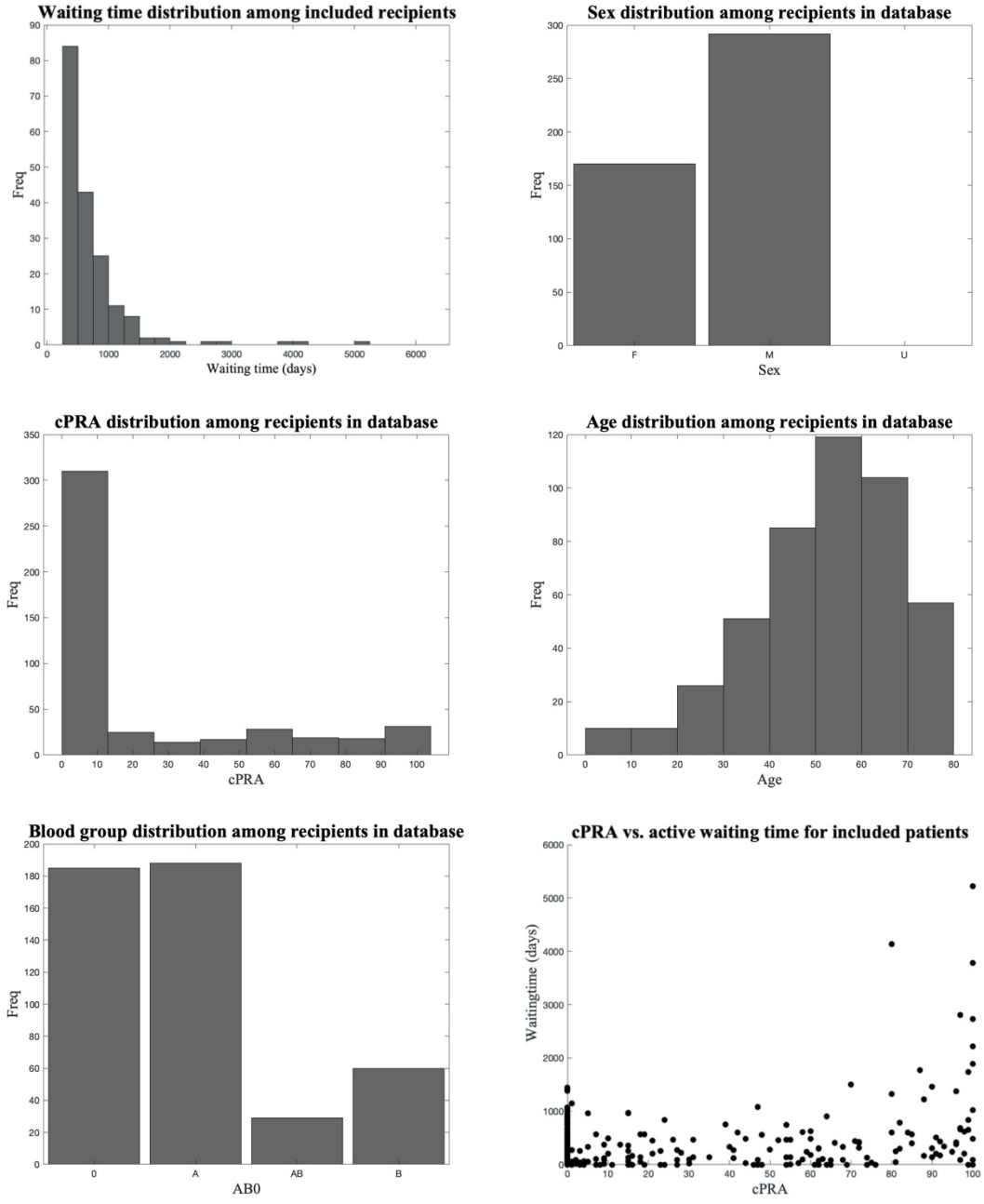


Figure 11: Analysis of donor dsata in the project database

Patients/Recipients



Base Figure 12: Analysis av patient data in the project database.

Base Case (n=150)

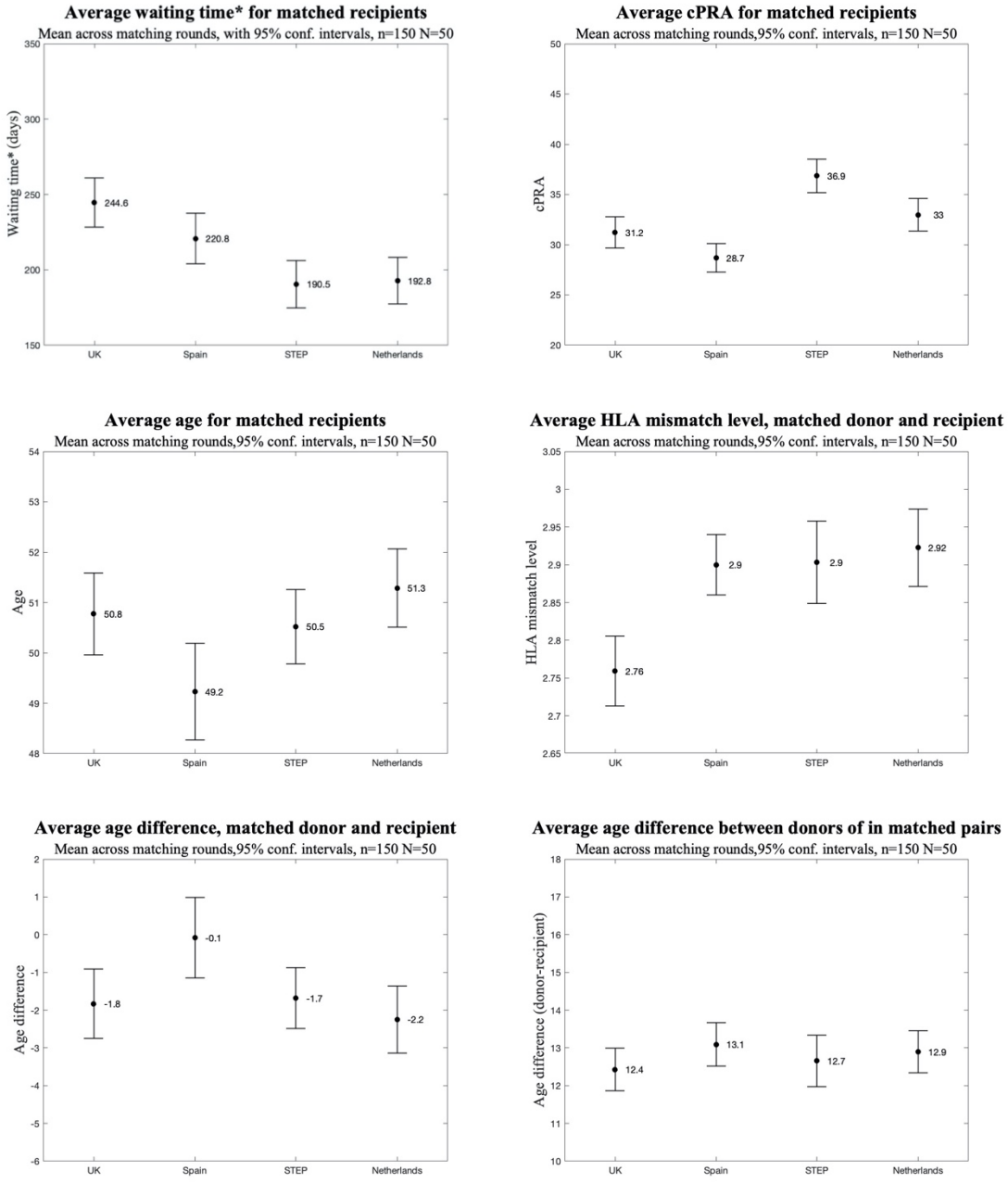


Figure 13:Means and error bars with 95% confidence intervals. UK left, Spain m.left, Scandinavia (STEP) m.right, Netherlands right

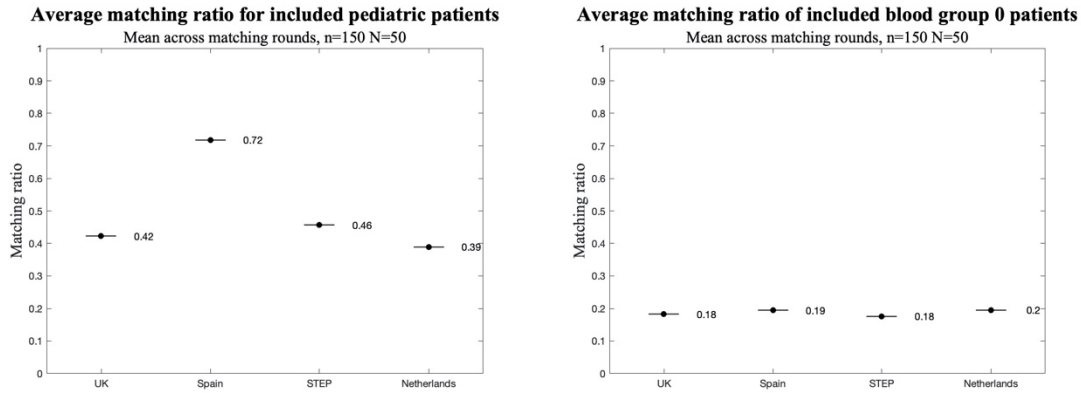


Figure 14: Plots for the base case. Means (center marker). UK left, Spain m.left, Scandinavia (STEP) m.right, Netherlands right

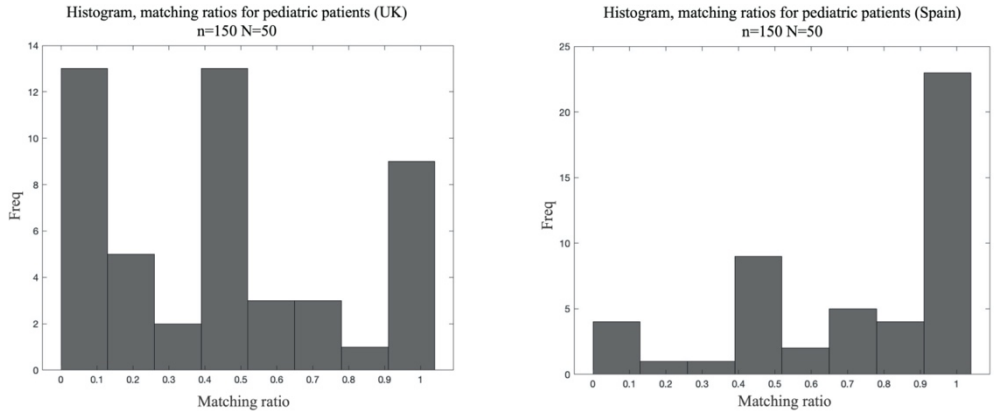


Figure 15: Histograms of matching ratios for pediatric patients for every matching round. UK left and Spain right

Increased number of pairs (n=250)

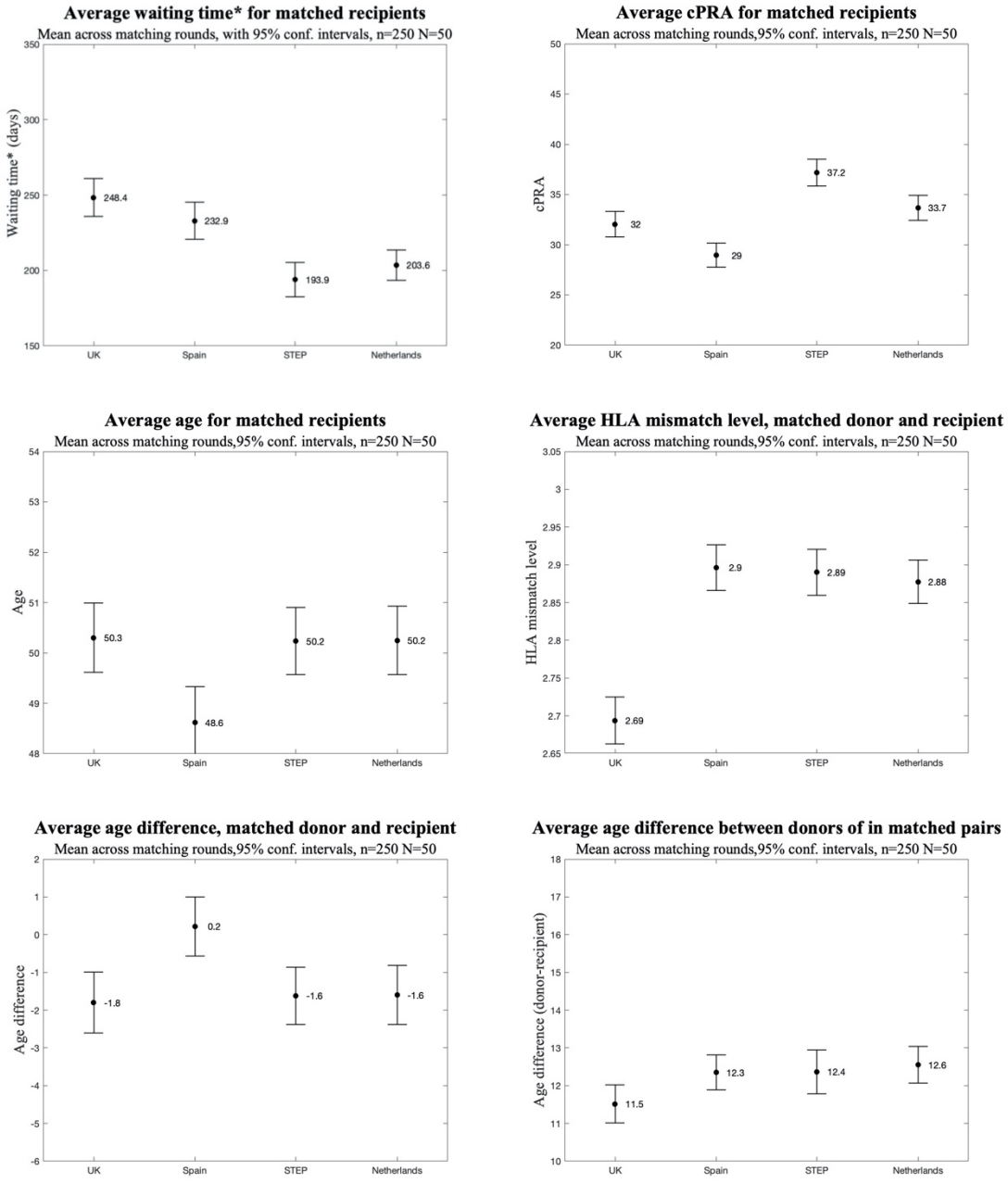


Figure 16: Means and error bars with 95% confidence intervals for the increased number case. UK left, Spain m.left, Scandinavia (STEP) m.right, Netherlands right

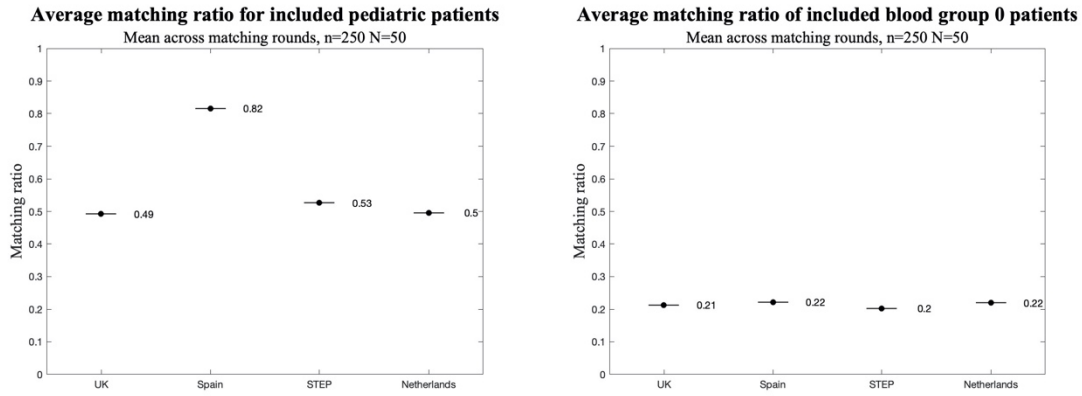


Figure 17: Means for matching ratios for the increased number case. UK left, Spain m.left, Scandinavia (STEP) m.right, Netherlands right

Decreased number of pairs (n=50)

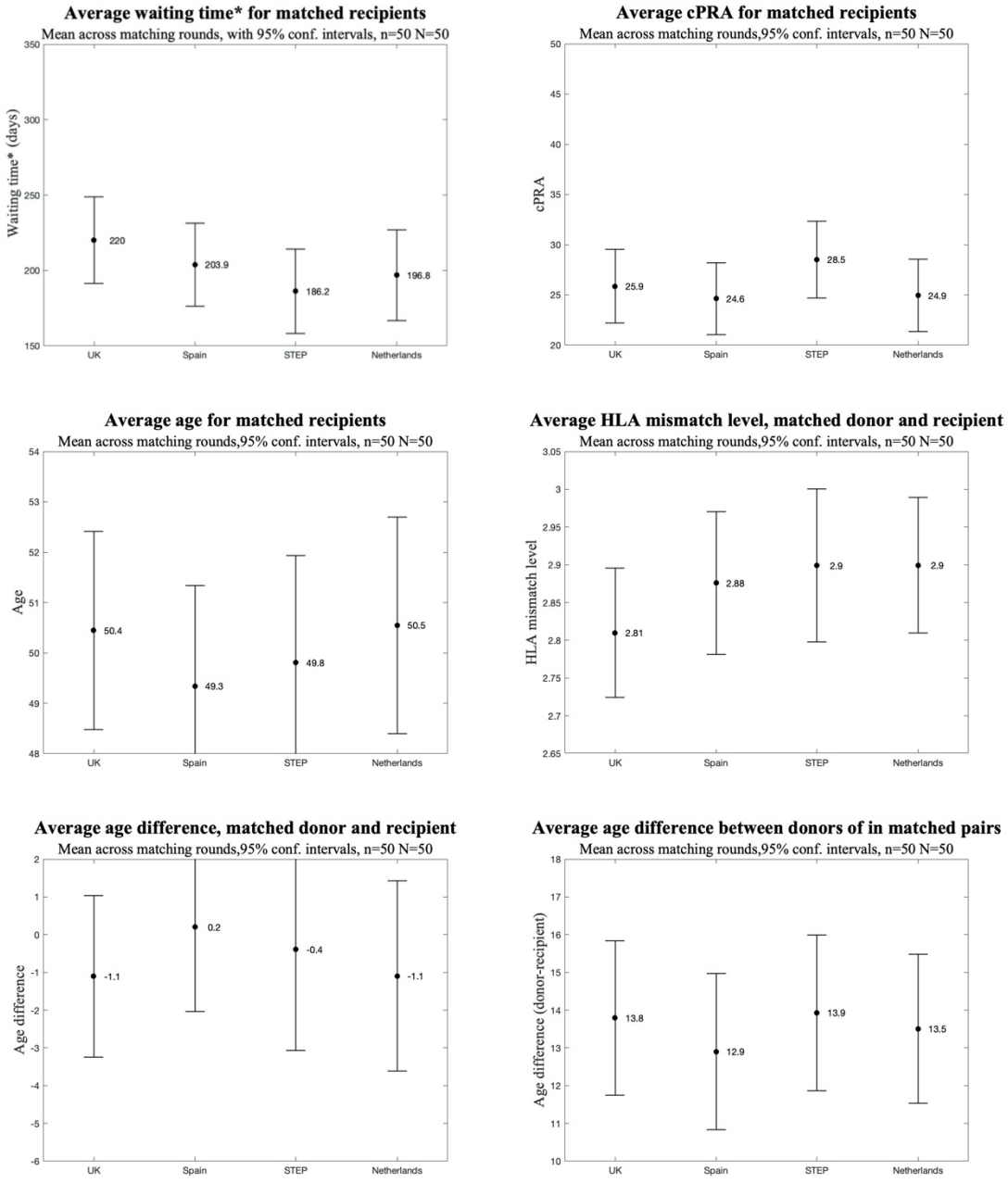


Figure 18:Means and error bars with 95% confidence intervals for the decreased number case. UK left, Spain m.left, Scandinavia (STEP) m.right, Netherlands right

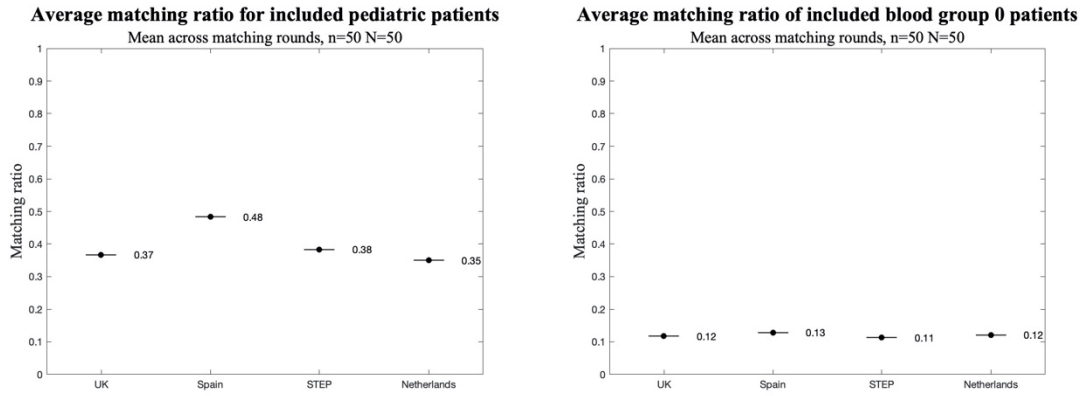


Figure 19: Means for matching ratios for the decreased number case. UK left, Spain m.left, Scandinavia (STEP) m.right, Netherlands right

Increased cPRA in project database (n=150)

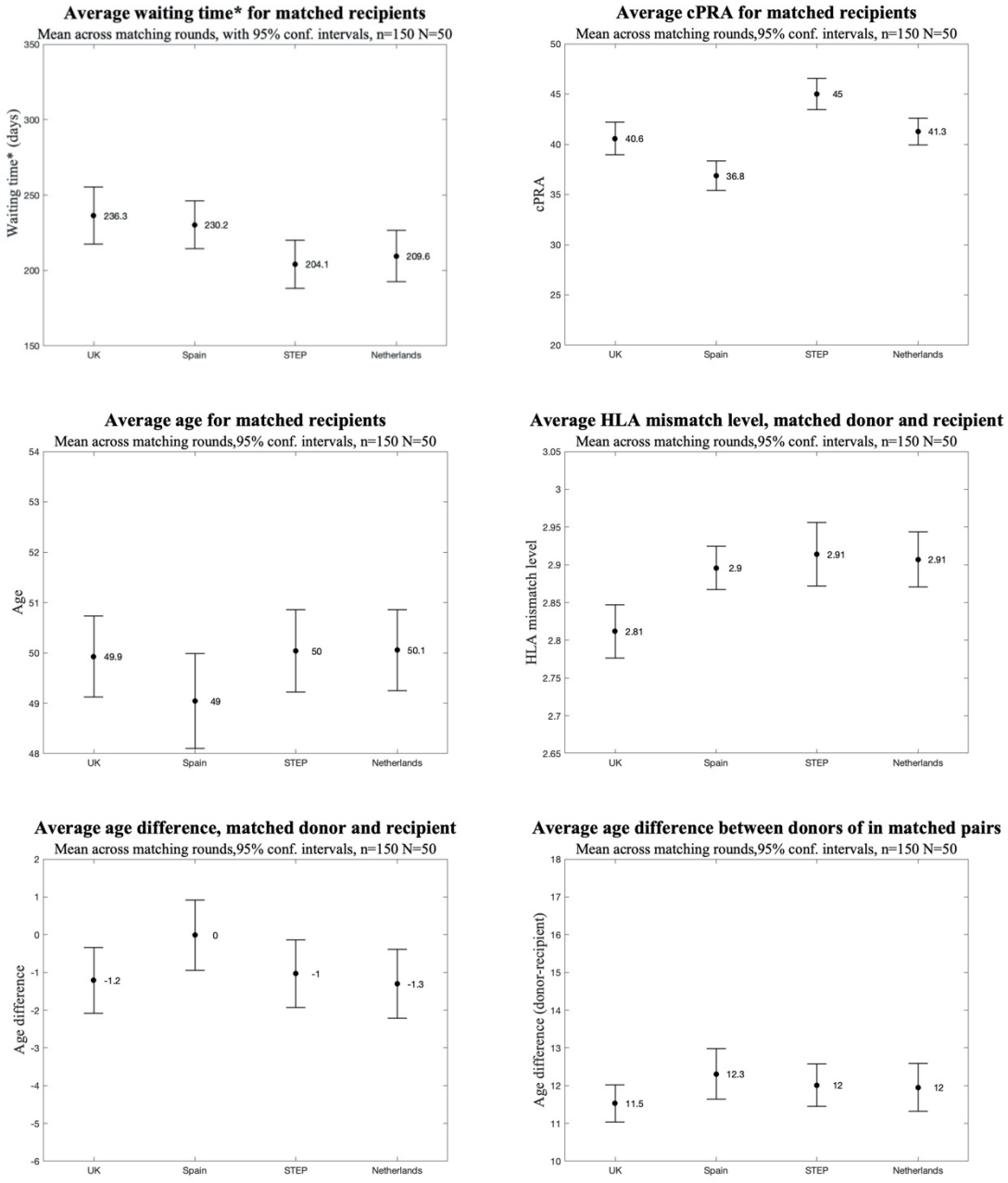


Figure 20:Means and error bars with 95% confidence intervals for the increased cPRA case. UK left, Spain m.left, Scandinavia (STEP) m.right, Netherlands right

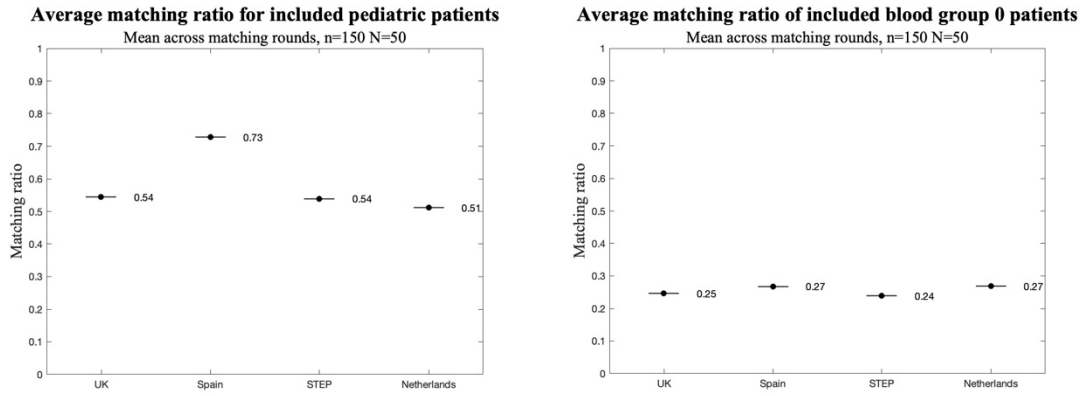


Figure 21: Means for matching ratios in the increased cPRA case

Appendix 2

Specification of the KEPs and Implementation in the Simulations

Below follows a specification of the priority rules in the KEPs, together with some additional comments regarding how they were implemented in our study. The rules used in our study are marked in bold, whilst the excluded ones are put in brackets.

Scandinavia (STEP)

The Scandinavian KEP, STEP is organized by the clinics and Scandiatransplant. The specification is taken from the Scandiatransplant website (29). In STEP *half compatible* pairs can join the program. A recipient with an incompatible donor can be matched with half and fully compatible donors but will strictly prefer the latter. A recipient with a half-compatible donor can enter the program, but will only be matched with fully compatible donors, as her or his own donors is preferred to any other half compatible donor. Altruistic donation is not allowed. 3-way exchanges are allowed. The hierarchical objectives are defined as follows:

1) **Maximum number of transplants**
2) **Priority for *hard-to-match* patients**, Measured by *match probability* (MP).
    a. This is defined in equation (1), similar to the definition in Keizer et al 2005 (35).

$$MP_i = (1 - PRA_i) * \frac{compatible\ donors\ in\ pool}{total\ donors\ in\ pool} \quad (1)$$

3) (Minimize number of transplants over blood group barrier)
    a. Not used here due to only AB0-compatbible transplants

Spain

The *Spanish National Kidney Exchange Program* (SNKEP), is organized by the clinics and *Organización Nacional de Transplantes* (ONT). The specification below was supplied by Elisabeth Coll, head of Medical Department at ONT(36). The Spanish program allows for 3-way exchanges, altruistic donations as well as for multiple donors being registered per recipient. Even compatible pairs are allowed to join the program with incentive to find a clinically better donor.. The hierarchical objectives are defined as follows:

1) **Maximum number of transplants**
2) A score:

- **Blood group**: 30 points for matchings withing the same blood group, with 30 additional if both are blood group 0
- **Hard-to match:** Points for high match probability MP. MP defined as in Keizer et al 2005 (35)

$$MP_i = (1 - PRA_i) * \frac{No.\,compatible\ donors}{No.\ AB0 compatible\ donors} \quad (1)$$

$$\begin{cases} MP\ [0\ 0.25] \rightarrow 30\ points \\ MP\ [0.26\ 0.50] \rightarrow 20\ points \\ MP[0.51\ 0.75] \rightarrow 10\ points \\ MP\ [0.76\ 1] \rightarrow 0\ points \end{cases}$$

- **Age difference/Mixed point**: Between 0 and 15 points. If any of the following apply point will automatically be 15:
    - Donor is younger than recipient
    - cPRA more or equal to 50%
    - Group 0 recipient or group AB donor (the recipient's own donor)
    - Recipient <18 years

  If none of the above criteria is met and donor-recipient age difference is over 20 years 0 points will be given and if difference under 11 years 15 points is given. If age difference is between 11 and 20, point is given between 0 and 15 according to
    - 11-11.99 years: 14 points
    - 12-12.99 years: 13 points
    - 13-13.99 years old: 12 points
    - 14-14.99 years old: 11 points
    - 15-15.99 years old: 10 points
    - 16-16.99 years old: 5 points
    - or 17-17.99 years 4 points
    - 18-18.99 years old: 3 points
    - 19-19.99 years old: 2 points
    - ≥20 years: 0 points
- **Time in program**: 30 points for more than 1 year
- **Child recipient**: 500 points for recipient under 18 years of age. Only donors aged under 50 will apply.
- (Time on dialysis: 0.05 points per month.)
    - *Not used here due to lack of data*

- (Compatible pairs: Patients with participating compatible donor will be awarded 250 points)
  - *Not used. Only incompatible were assumed to be included*
- (Desensitized patients: Patients that has gone through desensitization treatment will get 100 points.)
  - *Not used here due to lack of data*

United Kingdom

The UK Living Kidney Sharing Schemes (UKLKSS) is organized by National Health Service Blood and Transplant (NHSBT). The following specification is based on NHSBT documents for living donor kidney transplant.(28). Altruistic donation and 3-way cycles are allowed. The rules for AB0i-transplants are unclear and NHSBT has been contacted without receiving an answer.

. The hierarchical objectives are

1) (Maximize the number of effective 2-way exchanges)
   a. *Not used here due to only 2-way exchanges*
2) **Maximize total number of transplants**
3) (Minimize the number of 3-way exchanges)
4) (Maximize the number if embedded 2-way exchanges)

The score for the UKLKSS is as follows:

- **Previous matching rounds**. 50 points per number of quarterly matching rounds participated in without matching.
  - *Implemented using waiting time with 90 days/quarter*
- **Sensitization points**. Defined as

$$SP = \frac{cRF}{2}\ (\%)$$

  - *cRF assumed to be equal to cPRA*
- **HLA mismatch points**.0-15 points for mismatch level in HLA-DR and HLA-B
  - $$\begin{cases} Level\ 1\ 000 = 15\ points \\ Level\ 2\ [0\ DR\ and\ 0\ or\ 1\ B] = 10\ points \\ Level\ 3\ [0\ DR\ and\ 2\ B]\ or\ [1\ DR\ and\ 0\ or\ 1\ B] = 5\ points \\ Level\ 4\ [1\ DR\ and\ 2\ B] or\ [2\ DR] = 0\ points \end{cases}$$
- **Donor-Donor age difference**. 3 points if donor-donor age difference is less than 20 years.

Netherlands

The KEP in the Netherlands is organized by Nederlandse Transplantatie Veriniging (Dutch Transplant Foundation). The specification is taken from the ENCKEP handbook 20199 (21). Altruistic donation is allowed, and the maximum cycle length is 4. The Dutch algorithm relies solely on hierarchies as follows:

1) **Maximize the number of transplants**
2) **Maximize the number of blood type identical transplants**
3) **Match the patients in priority order based on match probability MP**
    a. $MP_i = (1 - PRA_i) * \frac{No.compatible\ donors}{No.\ AB0compatible\ donors}$

4) (Minimize the length of the longest cycle)
    a. *Not used due to only using 2-way exchanges*

2) (Maximize the spread over transplant centers per cycle and chain)
    a. *Not used due to data not applicable*

3) **Match the recipient with the longest waiting time.**