DiSIT, Computer Science Institute
Università del Piemonte Orientale "A. Avogadro"
Viale Teresa Michel 11, 15121 Alessandria
http://www.di.unipmn.it

**A dynamic simulation model for comparing kidney exchange policies**

*M. Beccuti, G. Franceschinis, S. Villa (beccuti@di.unito.it, giuliana.franceschinis@uniupo.it, silvia.villa@polimi.it)*

TECHNICAL REPORT TR-INF-2017-03-01-UNIPMN
*(March 2017)*

Research Technical Reports published by DiSIT, Computer Science Institute, Università del Piemonte Orientale are available via WWW at URL `http://www.di.unipmn.it/`.
Plain-text abstracts organized by year are available in the directory

## Recent Titles from the TR-INF-UNIPMN Technical Report Series

2016-04 *Tracing sharing in an imperative pure calculus*, P. Giannini, M. Servetto, E. Zucca, December 2016.

2016-03 *SUPPORTING DATA COMMUNICATION AND PATIENT ASSESSMENT DURING EMERGENCY TRANSPORTATION*, M. Canonico, S. Montani, M. Striani, September 2016.

2016-02 *TECHNICAL NOTE TO Forensic Analysis of the ChatSecure Instant Messaging Application on Android Smartphones (see below for citation details)*, C. Anglano, M. Canonico, M. Guazzone, September 2016.

2016-01 *Reasoning in a rational extension of SROEL*, L. Giordano, D. Theseider Dupré, May 2016.

2014-02 *A Provenly Correct Compilation of Functional Languages into Scripting Languages*, P. Giannini, A. Shaqiri, December 2014.

2014-01 *An Intelligent Swarm of Markovian Agents*, A. Bobbio, D. Bruneo, D. Cerotti, M. Gribaudo, M. Scarpa, June 2014.

2013-01 *Minimum pattern length for short spaced seeds based on linear rulers (revised)*, L. Egidi, G. Manzini, July 2013.

2012-04 *An intensional approach for periodic data in relational databases*, A. Bottrighi, A. Sattar, B. Stantic, P. Terenziani, December 2012.

2012-03 *Minimum pattern length for short spaced seeds based on linear rulers*, L. Egidi, G. Manzini, April 2012.

2012-02 *Exploiting VM Migration for the Automated Power and Performance Management of Green Cloud Computing Systems*, C. Anglano, M. Canonico, M. Guazzone, April 2012.

2012-01 *Trace retrieval and clustering for business process monitoring*, G. Leonardi, S. Montani, March 2012.

2011-04 *Achieving completeness in bounded model checking of action theories in ASP*, L. Giordano, A. Martelli, D. Theseider Dupré, December 2011.

2011-03 *SAN models of a benchmark on dynamic reliability*, D. Codetta Raiteri, December 2011.

2011-02 *A new symbolic approach for network reliability analysis* , M. Beccuti, S. Donatelli, G. Franceschinis, R. Terruggia, June 2011.

2011-01 *Spaced Seeds Design Using Perfect Rulers*, L. Egidi, G. Manzini, June 2011.

2010-04 *ARPHA: an FDIR architecture for Autonomous Spacecrafts based on Dynamic Probabilistic Graphical Models*, D. Codetta Raiteri, L. Portinale, December 2010.

# A dynamic simulation model for comparing kidney exchange policies

**Marco Beccuti**
*Dipartimento di Informatica, Università degli Studi di Torino*
*Corso Svizzera 185 - 10149 Torino, Italy*
*email: beccuti@di.unito.it*

**Giuliana Franceschinis**
*Università degli Studi del Piemonte Orientale, Dipartimento di Scienze e Innovazione Tecnologica*
*Viale Teresa Michel 11 - 15121 Alessandria, Italy*
*email: giuliana.franceschinis@uniupo.it*

**Silvia Villa**
*Dipartimento di Matematica, Politecnico di Milano*
*Piazza Leonardo da Vinci 32 - 20133 Milano, Italy*
*email: silvia.villa@polimi.it*

**Abstract**

In this paper we tackle the *dynamic* kidney exchange problem. First, we propose a formal definition through a mathematical model whose solution provides an optimal exchange policy, then we analyze two *almost myopic kidney exchange policies* which are currently adopted in practice by many transplant organizations. With respect to previous works in this area the models proposed in this paper take into account the quality of the performed exchanges.

The *dynamic* kidney exchange problem is formulated through a Markov Decision Process (MDP) model which allows to evaluate the performance of an optimal policy in a simplified case (constraining the state space size of the MDP). The results obtained through numerical solution of the MDP are then compared with the performance of two almost myopic policies evaluated using simulation. The obtained results show that in this case the performance of the considered policies is quite similar to that of the optimal policy.

The analysis is then extended to more realistic settings by means of a parametric simulator (freely available for download) that allows to consider multiple characteristics of both donors and recipients, different degrees of compatibility quantifying the quality of the exchange, and different cost factors concurring to the definition of the reward function on which the evaluation is based. The influence of the various parameters on the achievable reward is thoroughly discussed.

**Keywords**: Kidney exchange problem, Markov Decision Processes, Simulation

## 1 Introduction

Kidney transplantation is the elective treatment for many kidney diseases and it is nowadays a routine therapy. A patient needing a transplant may have a living donor, who is typically a friend, or a relative. However, in several cases, the donor is not compatible with the patient due to blood type or other physical characteristics, such as tissue type. To overcome this problem, and to increase the number of living donors, the possibility of organizing kidney exchanges has been suggested by the medical community in [45] and further investigated by many points of view in [33, 47, 57]. Nowadays, in many countries, incompatible pairs are allowed to swap donors when a mutual compatibility condition is satisfied. These swaps can involve more than two pairs, giving rise to cycles. The practical implementation of such exchanges raises several questions, and some of them have been faced using mathematical models. In particular, there is a rich literature devoted to the modeling and the efficient organization of kidney exchanges in the static setting [12, 13, 39, 55, 48, 49, 50, 51, 40]. In particular, an integer-programming model has been proposed in [15]. However, the real decision problem is better represented by a dynamic model. Indeed, donor-recipient pairs join the system over time, not all simultaneously. Most importantly, the matching policy, which regulates the exchanges, plays a key role in the efficiency of the kidney exchange system. This is why, starting from [54], the interest in the theoretical and practical aspects of dynamic matching mechanisms is growing [27].

In this paper, we consider a system where a centralized authority decides how to organize 2-way or 3-way exchanges over time, observing the evolution of the pool. More precisely, on the basis of the available information, i.e. the pairs who joined the system until the decision point, the decision maker chooses a feasible, possibly empty, set of exchanges (matching). This choice influences the composition of the pool for future exchanges, and has also an impact on the overall social welfare function employed by the institution in charge of the organization. Therefore, when taking a decision, the central authority has to consider the immediate costs and rewards, and also the future consequences of its decision.

The first contribution of this paper is a general model for the dynamic kidney exchange problem, showing that it can be conveniently formulated as a Markov Decision Process (MDP)[44] under reasonable assumptions. At each decision point the decision maker has to choose a matching from an available set of admissible exchanges. Moreover, a detailed probabilistic information describing the distribution of blood types, age, etc. of expected future arrivals is available, and should be taken into account. The performance of a chosen policy (i.e. a sequence of decisions) is measured taking into consideration the number and the quality of the executed transplants, the costs due to the organization of the exchanges and to unmatched patients still waiting. This reward function takes into account several aspects that are relevant in the definition of a kidney exchange problem. The definition of the reward could be thought as the scalarization of a multi-criteria optimization problem, where the objectives to optimize include for instance the total waiting time for each pair, the quality and the expected number of matchings. Since such MDP has a high dimension in terms of possible states and available decisions, known optimization techniques are not capable to solve it efficiently, and only heuristics have been considered so far to approximate its solution (see [7]). In realistic situations it is thus practically impossible to obtain an optimal policy by solving the MDP, but it is interesting to evaluate the performance of policies that are adopted in practice using the MDP as reference model for comparison on small size problems.

Among the possible dynamic matching policies, in this paper we focus on two specific classes of policies adopted in practice by many transplant organizations all over the world (see for instance [20, 32, 33]): they consist in a sequence of maximum weight matchings, i.e. sets of 2-way and 3-way exchanges maximizing a suitably chosen welfare function in the static situation. The first class organizes maximum weight matchings at regular intervals of time, while the second performs them only when the number of waiting pairs in a certain class is above a given threshold. We call these policies *almost myopic*, because they provide two possible decisions: maximizing the immediate reward (as a *myopic* policy), or just waiting until the next decision point without doing any exchange.

The second contribution presented in this paper is a parametric simulator used to study these almost myopic policies under various conditions. The simulation results on the two considered classes of almost myopic policies are first compared with the optimal exchange policy derived by the MDP numerical solution on a stylized but significant example: this example represents a simple model of the real situation, in which the types are defined considering the blood type of the recipient and of the donor, and by the age of the donor (young or mature). We consider only pairs in which the donor is of type B and the recipient is of type A and vice versa. In this stylized model, the relevant features of the entire real situation are retained: the compatibility constraints are present, and the difference in the quality of the transplants is determined by the age of the donors (see also [40]). From the analysis of this case study we observe that the performance (in terms of transplant quality) of the considered policies, evaluated through simulation, is quite similar to that of the optimal policy computer from the MDP. Our simulation study is then extended to more realistic settings by means of an ad-hoc developed parametric simulator that allows us to consider multiple characteristics of both donors and recipients, different degrees of compatibility (quantifying the quality of the transplant) and different cost factors concurring to the definition of the reward function on which the evaluation is based. The influence of the various parameters on the achievable reward is investigated in the paper.

The main original contribution of the work presented in this paper is thus twofold: a general model able to take into account the quality of the transplants and other factors such as organizational costs, and the implementation of a parametric simulator for the evaluation of two classes of policies that can be applied to the model. Indeed, the quality of the transplants is an important aspect; since the relationship between the graft survival and some characteristics of the donor and the recipient has been highlighted by many papers (see [25, 38, 41, 42]) in the medical literature. The simulator is available free of charge for academic institutions and non profit organizations; it can run on Linux, Unix and MacOS X system. The ultimate goal of the results presented in this paper is not to give recommendations in terms of the best policy to be used, but rather the proposal of a flexible and easy to use simulation tool to investigate the effect of different choices of the parameters.

The paper is organized as follows: in Section 2 the static problem is introduced, mainly focusing on the aspects which are useful in the subsequent description of the dynamic case. In Sections 3 and 4 the proposed MDP model for the dynamic case and the almost myopic policies are described. Section 5 reports a comparison between almost myopic policies and the optimal exchange policy derived by the MDP considering a simplified case study. In Sections 6 and 7 our parametric simulator is thoroughly described and used to analyze almost myopic policies in a more realistic way. Finally, Section 8 reports some concluding remarks and future works.

## 1.1 Related work

The first model of a dynamical kidney exchange problem has been considered in [54], where a compatibility of type 0-1 between patients and donors is assumed. A characterization of the optimal policy is given when the model depends only on the blood type, and the costs are due to patients in the waiting list. In this setting it is proved that in the case of 2-way exchanges the optimal policy is the one which organizes exchanges as soon as they become available. In the same paper it is also proved that the same conclusion does not hold if longer cycles are admitted. In [7] the probabilistic information about the characteristics of patients is deeply exploited and several stochastic and online learning algorithms are presented to approximate an optimal policy, which provide encouraging results on simulated and real data. In [6], the authors focus on tissue type incompatibility, while abstracting away from blood-type compatibility, and the considered matching policies are similar to the ones studied in this paper. In their setting, the authors show the importance of allowing chains initiated by an altruistic donor to increase the number of possible matchings, as it was observed also in [21]. Another matching policy, based on weights assigned to different matches is proposed in [22]. Another recent line of work takes into account the possibility of match failure when maximizing the expected number of transplants, see [23]. The problem of including additional constraints in order to ensure fairness of the proposed matching policy is studied in [24]. A different dynamical model based on stochastic games is considered in [34], where the fact that the health status of the patient worsens along time is taken into account. A slightly different problem is considered in [58], where the effects of the exchange system are studied in relationship with the waiting list for organs from deceased donors.

## 2 The static model

In this section we are going to present a model for the static kidney exchange problem, where a set of donor-recipient pairs is given, and the best (in some sense) admissible set of 2-way (or 2-way and 3-way) exchanges has to be selected. The restriction to 2-way and 3-way exchanges derives from practical constraints, since in the real world the transplants involved in an exchange must be performed at the same time. As already mentioned in the introduction, this problem has been extensively studied (e.g. [12, 39, 48, 49, 50, 51, 55]). Given a set of donor-recipient pairs available for the exchanges, we introduce a model which takes into account the qualitative properties of the involved transplants. To this aim we fix a certain number of variables, clinical or not, such as blood type, tissue type, age, body mass index and general health condition, as well as number of days spent on the waiting list by the patient, or place where the patient lives. It is not restrictive to model these characteristics as discrete variables ranging on a finite number of values. In fact, for some of the mentioned characteristics this is natural, while for others can be imposed by choosing a suitable discretization.

The same set of variables can be considered not only for the patient, but also for the donor. On the basis of the values assumed by the chosen variables (both for the patient and the donor), we can then categorize each donor-recipient pair into a finite set of pair types $\mathcal{T} := \{1, \dots, n\}$. The classification into distinct types allows us to incorporate in the model both the features which are relevant in determining the quality of the transplant, and information of different origin such as the geographic distance between pairs of different types. This possibility could be relevant in the real application of the model.

The quality of a $k$-way exchange, with $k \in \{2, 3\}$ is then measured in terms of a reward, taking into account the relevant features of such exchange. A natural reward is obtained summing up individual rewards measuring the quality of each involved transplant, but also other choices may be considered. The standing assumption throughout the paper is the fact that *the quality of the exchanges is completely determined by the types of the involved pairs.* We model this by introducing a $n \times n$ compatibility matrix $R$, where the entry $R_{ij}$ expresses the quality of a transplant from a donor of a pair of type $i$ to a recipient of a pair of type $j$. In particular, $R_{ij} = 0$ means that a transplant from a pair of type $i$ to one of type $j$ is not feasible.

Mathematically speaking a *(static) kidney exchange problem* consists of the following objects:

a) a set of pairs, which can be identified with $(s_1, \dots, s_n) \in \mathbb{N}^n$ describing the number of pairs of each type;

b) the compatibility matrix $R$ of size $n \times n$ defined above.

Starting from an $n$-tuple of non negative integers $\mathbf{s} = (s_1, \dots, s_n)$ and the compatibility matrix $R$, a kidney exchange problem can be identified with a directed weighted graph $G(\mathbf{s})$, with $s_1 + \ldots + s_n$ vertices, where a directed edge $(i, j)$ is present in the graph if and only if, called $t_i$ and $t_j$ the types of the pairs $i$ and $j$, we have $R_{t_i t_j} > 0$. The weight of each edge is $w_{ij} := R_{t_i t_j}$.

A major aim of the literature on static kidney exchange problems is to devise a mechanism, i.e. a procedure assigning a matching to each kidney exchange problem satisfying specific properties.

**Definition 1.** ***(Matching)*** *Let $V$ be a finite set, let $E \subset V^2$, and let $v : E \to [0, +\infty[$. Let $G = (V, E, v)$ be the corresponding directed weighted graph. A matching in $G$ is a subset of cycles (possibly of length less than or equal to a given limit $K$) without nodes in common.*

In the above definition vertices represent donor-recipient pairs, while weighted and directed arcs represent the compatibility of the donor in the source node and the recipient of the destination node. A cycle involving $k$ vertices represents a $k - way$ exchange. The point of view is the one of a central authority that must choose which exchanges to perform. If no distinction between compatibility levels is made (i.e. if all graph arc weights are equal), in [49] the matching guaranteeing a transplant to the highest possible number of participants (maximum cardinality matching) is proposed as a solution. A solution taking into account the quality of the exchanges is proposed in [48], but without restrictions on the cycle length, i.e. on the length of the number of pairs involved in the same chain of exchanges. Defining $\mathcal{M}(\mathbf{s})$ the set of matchings in $G(\mathbf{s})$, and given an objective function $f : \mathcal{M}(\mathbf{s}) \to [0, +\infty)$, measuring the quality of a matching, in [55] a matching maximizing $f$ was proposed as a solution. In this paper, we will consider matchings involving 3 pairs at maximum, i.e. we will restrict to matchings involving only 2-way and 3-way exchanges. Thus, from now on we will call matching a matching where the length of the cycles is less or equal to 3. Given a function $f$ we will use the following notation:

$$OM(G(\mathbf{s})) := \{\text{matchings involving 2-way and 3-way exchanges in } G(\mathbf{s}) \text{ maximizing } f\}. \quad (1)$$

where $G(\mathbf{s})$ is the graph whose vertices are all pairs present in state $\mathbf{s}$ and whose arcs connect all compatible pairs according to the compatibility matrix $R$. In particular, if $f(m) = \sum_{(i,j)\in m} w_{ij}$, $m \in \mathcal{M}(\mathbf{s})$ the solution is identified with a maximum weight matching, which can be computed efficiently, using e.g. the algorithm proposed in [1] (see also Section 6).

In the real world, a necessary condition in order to have compatibility between a donor and a recipient is the so called "negative cross-match". Cross-matching is a very sensitive and final test whose result is not totally predictable starting from physical characteristics as the ones we mentioned in the definition of the types, but should be verified case by case. To deal with the crossmatch it is possible to consider a probability of compatibility rather than a deterministic compatibility between types. In this paper we use a deterministic matrix, since including this type of possibility does not require different analysis techniques. Indeed, after a matching has been devised, it is enough to discard a subset of transplants, based on the realization of a random variable. Since we do not take into account the possibility of having a positive crossmatch when organizing the exchanges, the subsequent results overestimate the number of possible transplants.

This assumption, while simplifying the real situation, is considered acceptable also from a medical point of view and is common in the literature studying kidney exchange problems [49].

# 3 The dynamic model

As discussed in the introduction, the static model is not enough to correctly face the kidney exchange problem, which is intrinsically dynamic, since the pool of pairs waiting a transplant evolves over time. The solution proposed in the static setting is no longer optimal if dynamic effects are considered.

Indeed, inefficiency in a dynamic environment could arise if decisions are taken without considering the consequences of present decisions on future possibilities. In particular, the choice of the exchange timing is relevant. We now propose a more realistic dynamic model, assuming that each pair enters the process and leaves it only upon receiving a kidney. Moreover, we assume that the pairs of each type $i$ join the system according to a Poisson arrival process denoted by $N_i(t)$ and that the arrival processes of each type are independent. Let $\lambda_i$, $i \in \{1, \dots, n\}$ the arrival rates of each type. Such an assumption is considered consistent with the real situation and it is adopted in many papers treating organ allocation [46, 58, 54, 11]. By definition of Poisson process, the probability of having $h$ arrivals of pairs of type $i$ in the time unit, is independent from the specific time period and satisfies

$$\mathbb{P}[N_i(t) = h] = \frac{\lambda_i^h}{h!} e^{-\lambda_i}. \tag{2}$$

To deal with the pool's evolution, in this section we show that the problem of dynamic kidney exchange can be conveniently formulated as an MDP [30, 44]).

**Definition 2.** *A Markov decision process is a* 4*-tuple* $(\mathcal{S}, \mathcal{A}, r, p)$ *where*

- $\mathcal{S}$ *is a set of states;*
- $\mathcal{A} = (\mathcal{A}_\mathbf{s})_{\mathbf{s} \in \mathcal{S}}$*, where* $\mathcal{A}_\mathbf{s}$ *is the set of actions available in state* $\mathbf{s}$
- $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ *is the immediate reward function;* $r_\mathbf{s}(m)$ *is the reward determined by choosing action* $m$ *being in state* $\mathbf{s}$*;*
- $p$ *is a probability distribution on* $\mathcal{S} \times \mathcal{S} \times \mathcal{A}$*,* $p_{\mathbf{s}\mathbf{s}'}(m)$ *is the probability of passing from the state* $\mathbf{s}$ *to the state* $\mathbf{s}'$ *having chosen the action* $m$*.*

We now discuss the choices of the various elements in Definition 2 for the kidney exchange problem. Let $T$ denote the set of time points in which a decision can be taken: we will always assume that $T$ is discrete and infinite. Such problems are classified as infinite horizon problems. This seems a reasonable assumption, since the exchanges are likely to take place for a long period of time.

**State of the system.** The system at time $t$ is constituted by the set of donor-recipient pairs available for an exchange at time $t$. We assume that there is a maximum number of pairs of each type admitted to the system, and we call it $N$. Therefore, if $n$ denotes the total number of pair types, the state of the system is described by a vector $\boldsymbol{s} \in \{0, \dots, N\}^n$, where each component $s_i$ is the number of pairs of type $i$ present in the system at the considered time point. We denote the set of admissible states at time $t$ by $\mathcal{S}$, since it is independent of $t$.

**Set of actions.** When the decision maker observes the state $\boldsymbol{s}$, she/he chooses an action from $\mathcal{A}_{\boldsymbol{s}}$, the set of allowable actions in $\boldsymbol{s}$. If the observed state is $\boldsymbol{s}$, the available actions are the matchings (included the empty one) involving only 2 and 3-way exchanges, of the weighted graph constructed on $\boldsymbol{s}$ according to the compatibility matrix $R$, and denoted by $G(\boldsymbol{s})$ (see Section 2). More precisely:

$$\mathcal{A}_{\boldsymbol{s}} = \{a_m \,:\, a_m \text{ is a matching in } G(\boldsymbol{s})\} := \mathcal{M}(G(\boldsymbol{s})).$$

**Immediate rewards.** Given the state of the system and the chosen action, the immediate reward is obtained by subtracting some costs to the gains due to the performed transplants. There are some general fixed costs which are independent of whether the exchanges are organized or not, and other costs that are related to the exchange procedures. The general fixed costs can be neglected, since they are constant with respect to the adopted policies. We model the costs dividing them into two categories: the first one includes the costs due to the organization of the exchanges, while the second takes into account the unmatched patients still on the waiting list. More precisely, denoting by $N(\boldsymbol{s}, a_m)$ the number of transplants performed by choosing action $a_m$ in state $\boldsymbol{s}$, the total cost of a matching is

$$C(\boldsymbol{s}, a_m) = \begin{cases} 0 & \text{if } N(\boldsymbol{s}, a_m) = 0 \\ F + c_t N(\boldsymbol{s}, a_m) & \text{otherwise,} \end{cases} \tag{3}$$

where $c_t$ is the unitary cost of a transplant. The fixed costs $F$ may be interpreted as the costs which are independent from the number of matched patients, but appear in correspondence of any exchange round (management of the waiting list, special working staff, etc.). The variable costs are assumed to be proportional to the number of matched patients, and they are due for instance to the preliminary examinations, the surgeries, therapies, etc. Reasonable choices of these parameters are discussed in the following.

The second kind of costs derives from the patients that are waiting, and is linear in the number of patients that remain unmatched under action $a_m$, namely:

$$C_w(\boldsymbol{s}, a_m) = \left( \sum_{i=1}^{n} s_i - N(\boldsymbol{s}, a_m) \right) c_u,$$

where $c_u$ is the unitary cost of waiting for a time unit. Finally, for a given objective function $f$ measuring the matching quality (e.g. quality adjusted life years expected by the patient as result of an exchange), the welfare function can be defined as follows:

$$r_{\boldsymbol{s}}(a_m) = f(a_m) - C(\boldsymbol{s}, a_m) - C_w(\boldsymbol{s}, a_m).$$

**Transition probabilities.** It follows from our assumptions that the state of the system at time $t$ depends only on the observed state at time $t-1$, the arrivals and the chosen matching and does not depend on the history of the process (due to the Markov property). The state at the subsequent decision point is therefore entirely determined by a transition law. Thanks to our assumptions about the arrival processes (see equation 2), the transition probabilities can be explicitly computed and are stationary, i.e. they do not depend on time.

Denoting by $\boldsymbol{m} = (m_1, \ldots, m_n)$ the number of matched pairs of each type, and assuming that a pair leaves the pool once it is matched and its recipient receives the selected transplant immediately, we get

$$p_{\boldsymbol{s}\boldsymbol{s}'}(\boldsymbol{m}) = \begin{cases} 0 & \text{if } \exists\, i \text{ s.t. } s'_i < s_i - m_i \\ \prod_{i=1}^{n} \mathbb{P}[N_i = s'_i - (s_i - m_i)] & \text{otherwise.} \end{cases} \tag{4}$$

**Policies.** The objective of the central authority organizing the exchanges is to find an optimal policy according to certain criteria, as we will explain later. A policy $\pi = (\pi^t)_{t \in T}$ (where $T$ is the set of decision points) is a sequence of decision rules, where $\pi^t$ in general may depend on the past history of the system. More formally, we define the set of histories of the system up to time $t$ as

$$\begin{aligned} H_t \;=\; & \{(\boldsymbol{s}_0, a_0, \ldots, \boldsymbol{s}_{t-1}, a_{t-1}, \boldsymbol{s}_t) \,:\, \boldsymbol{s}_k \in \mathcal{S}, \text{ for} \\ & k = 1, \ldots, t;\; a_k \in \mathcal{A}_{\boldsymbol{s}_k} \text{ for } k = 1, \ldots, t-1\}. \end{aligned}$$

A (deterministic) decision rule $\pi^t : H_t \to A_{\boldsymbol{s}_t}$ at time $t$ specifies the action chosen with regard to each history.

If the decision rule $\pi^t$ is independent of $((\boldsymbol{s}_0, a_0), \dots, (\boldsymbol{s}_{t-1}, a_{t-1}))$ for every $t \in \mathbb{N}^+$ the policy is said to be a Markov policy. If moreover the decision rules are independent of the time point $t$, i.e. $\pi^0 = \pi^1 = \dots$ the policy is called stationary. A stationary policy can be seen as a function from $\mathcal{S}$ to $\cup_{\boldsymbol{s} \in \mathcal{S}} \mathcal{A}_{\boldsymbol{s}}$, with $\pi(\boldsymbol{s}) \in \mathcal{A}_{\boldsymbol{s}}$ for every $\boldsymbol{s} \in \mathcal{S}$.

A matrix $P = (p_{ij})_{i,j=1,\dots,n}$ is said to be a transition matrix if $p_{ij} \geq 0$ for all $(i,j)$ and $\sum_{j=1}^{n} p_{ij} = 1$ for all $i = 1, \dots, n$. Each Markov policy induces a transition matrix $P(\pi^t)$ and a reward vector $r(\pi^t)$, defined as follows:

$$\begin{array}{ll} P(\pi^t)_{\boldsymbol{s}\boldsymbol{s}'} = p_{\boldsymbol{s}\boldsymbol{s}'}(\pi^t(\boldsymbol{s})) & \forall\, \boldsymbol{s}, \boldsymbol{s}' \in \mathcal{S} \times \mathcal{S} \text{ and } t \in T; \\ r(\pi^t)_{\boldsymbol{s}} = r_{\boldsymbol{s}}(\pi^t(\boldsymbol{s})) & \forall\, \boldsymbol{s} \in \mathcal{S} \text{ and } t \in T. \end{array}$$

**Optimality criterion.** Several optimality criteria exist for infinite horizon problems. This paper focuses on the maximization of the total expected $\delta$-discounted reward, which is defined as follows for a Markov policy $\pi$ when the initial state of the system is $\boldsymbol{s}_0$:

$$v^{\delta}_{s_0}(\pi) = \sum_{t=0}^{+\infty} \delta^t \sum_{\mathbf{s} \in \mathcal{S}} \mathbb{P}_{\boldsymbol{s}_0, \pi}[X_t = \boldsymbol{s}] r_{\boldsymbol{s}}(\pi^t(\boldsymbol{s})), \tag{5}$$

where $X_t$ is the random variable describing the state of system at time $t$, and $\mathbb{P}_{\boldsymbol{s}_0,\pi}[X_t = \boldsymbol{s}]$ denotes the probability that at time $t$ the state is $\boldsymbol{s}$, given that policy $\pi$ is used and the initial state is $\boldsymbol{s}_0$. More explicitly we can write:

$$\boldsymbol{v}^{\delta}(\pi) = \sum_{t=0}^{+\infty} \delta^t \boldsymbol{P}(\pi^0) \cdots \boldsymbol{P}(\pi^t) \boldsymbol{r}(\pi^t),$$

which for a stationary policy becomes

$$\boldsymbol{v}^{\delta}(\pi) = \sum_{t=0}^{+\infty} \delta^t \boldsymbol{P}(\pi)^t \boldsymbol{r}(\pi).$$

**Definition 3. Optimal policy.** *A policy $\pi_*$ is optimal if*

$$v^{\delta}_{\boldsymbol{s}}(\pi_*) \geq v^{\delta}_{\boldsymbol{s}}(\pi) \qquad \forall\, \boldsymbol{s} \in \mathcal{S} \quad \textit{and} \quad \forall \pi \in \mathcal{P},$$

*where $\mathcal{P}$ denotes the set of admissible policies.*
*We define $\mathbf{v}^{\delta} = (v^{\delta}_{\mathbf{s}}(\pi_*))_{\mathbf{s} \in \mathcal{S}}$ the optimal reward associated with the optimal policy $\pi_*$.*

The following theorem shows that for MDPs it is not restrictive to consider only stationary policies, since an optimal policy of this type can always be found.

**Theorem 1.** *Assume that $\mathcal{S}$ is finite, and let $\mathbf{v}^{\delta}$ be the optimal reward as defined in Definition 3. Then $\mathbf{v}^{\delta}$ is the unique fixed point of the function*

$$\begin{array}{rccl} U: & \mathbb{R}^{|\mathcal{S}|} & \to & \mathbb{R}^{|\mathcal{S}|} \\ & (Ux)_{\mathbf{s}} & = & \max_{a \in \mathcal{A}_{\boldsymbol{s}}} \{ r_{\mathbf{s}}(a) + \delta \sum_{\boldsymbol{s}'} p_{\boldsymbol{s}\boldsymbol{s}'}(a) x_{\boldsymbol{s}} \}. \end{array}$$

*Let $g_x$ be such that*

$$g_x(\boldsymbol{s}) \in \operatorname{argmax}_{a \in A_{\boldsymbol{s}}} \{ r_{\boldsymbol{s}}(a) + \delta \sum_{\boldsymbol{s}'} p_{\boldsymbol{s}\boldsymbol{s}'}(a) x_{\boldsymbol{s}} \}.$$

*Then the stationary policy defined by $\pi(\boldsymbol{s}) = g_{\mathbf{v}^{\delta}}(\boldsymbol{s})$ for all $\boldsymbol{s} \in \mathcal{S}$ is optimal.*

The proof of Theorem 1 follows directly from Theorems 6.2.3, 6.2.5 and 6.2.10 in [44]. The result can be directly applied to the kidney exchange problem in order to state that an optimal solution exists according to the optimality criterion given by the maximization of the expected total discounted reward.

The main drawback of this approach is that in practice the set of admissible states and of admissible actions is huge, growing more than exponentially both with respect to the number of types and with respect to the number of admitted pairs. This prevents the use of classical algorithms such as value iteration or policy iteration (see [44]). An alternative approach which avoids the complete computation of the state space and of the action space is discussed in [7]. The main idea is to find a sub-optimal policy by using only a small subset of the possible scenarios. To deal with the complexity of the model, in this paper we are going to propose two special sub-optimal policies.

# 4 Almost myopic policies

As a starting point of the dynamic analysis we focus on a specific class of policies, that we call almost myopic.

**Definition 4. (Almost myopic policy)** *A policy is called almost myopic if at each decision point the decision maker can choose between the empty matching and one maximizing the objective function $f$ (see* (1)*).*

We call these policies almost myopic because they are similar to the myopic policy, the main difference between the two is the possibility of choosing the empty matching (while myopic policies match compatible pairs as soon as they enter the system). We will consider two kinds of almost myopic policies: elements of the first kind schedule a matching maximizing the welfare function at regular intervals of time while policies of the second kind organize matching maximizing the welfare function only when the number of pairs in the pool is above a fixed threshold.

There is no evidence that an optimal policy should be of this form. Nonetheless, almost myopic policies exhibit some features that make them sufficiently interesting. In fact, among several existing programs of kidney exchanges around the world, there are at least two "common" procedures in organizing kidney exchanges: the first one consists in organizing the exchanges as soon as they become available (see [54]); the second one is to organize an exchange round at regular intervals of time, see [32, 33].

Given a dynamic kidney exchange problem, using the notations of Section 3 and recalling Equation 1, a time dependent almost myopic policy is defined as follows

$$\pi_\tau(\boldsymbol{s}, t) = \begin{cases} m^* \in OM(G(\boldsymbol{s})) & \text{if } t = k\tau \text{ for some } k \in \mathbb{N}, \\ \text{empty matching} & \text{otherwise,} \end{cases} \tag{6}$$

where $\tau$ is an a priori fixed time interval. If $OM(G(\boldsymbol{s}))$ contains more than one element, we choose one of them, trying to favor the ones involving types that are subject to a long waiting time (see Section 6 for more details). We call the time instants when a matching is organized an *exchange round*.

Analogously, we define a state dependent almost myopic policy as follows:

$$\xi_N(\boldsymbol{s}) = \begin{cases} m^* \in OM(G(\boldsymbol{s})) & \text{if } \sum_{i \in \mathcal{O}} s_i \geq N, \\ \text{empty matching} & \text{otherwise,} \end{cases} \tag{7}$$

where $\mathcal{O}$ is a subset of the pair types. We consider only a subset of all possible types since, as noted in [54], under a long-run assumption, there will be in the pool an accumulation of so-called *under-demanded* pairs. The *under-demanded* pairs can be identified from the medical data and the arrival rates, and are a subset of pairs for which incompatibility is due to blood-type.

# 5 Optimal and almost myopic policy comparison on a simplified model

In this section we estimate through simulation the total expected $\delta$-discounted reward as defined in (5), achieved by the policies $\pi_\tau$ and $\xi_N$ for different choices of $\tau$ and $N$, and compare it with the one corresponding to the optimal policy that we will obtain by solving the MDP. The goal of the comparison is twofold: on the one hand it is made between policies of the same class, in order to gain some information on the most appropriate choices for $\tau$ and $N$; on the other hand we aim at assessing the quality of these exchange policies with respect to the optimal one. In these simulations we will always consider as welfare function the one summing up the weights of the exchanges, so that the efficient matchings are the maximum weight ones.

In order to make the problem numerically tractable, we analyze a stylized model of the kidney exchange problem, restricting the number of possible pair types. Instead of considering all the possible blood types we focus our attention only on the pairs in which the donor is of type A and the intended recipient is of type B and the reciprocally compatible pairs in which the donor is of type B and the recipient of type A. As mentioned in the introduction, one of the novelties of our approach is the possibility of taking into account the quality of the exchanges. There are many factors influencing the graft survival, see e.g. [25, 31, 38, 32, 41, 42, 39] for an analysis of this issue. In this example we focus on the age of the donors, that we classify into two distinct classes: *young* or *mature* (see also [40]). We suppose therefore that there are only four pair types, and we call them 1 (recipient of type B and a young donor of type A), 2 (recipient of type A and a mature donor of type B), 3 (recipient of type B and a mature donor of type A) and 4 (recipient of type A and a young donor of type B). In agreement with the assumptions made in the description of the model, pairs of type $i$ ($i = 1, 2, 3, 4$) arrive according to a Poisson process with rate $\lambda_i$ and they all join the exchange system. Reciprocal compatibilities are described by the following compatibility table:

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0 | $l/2$ | 0 | $M/2$ |
| 2 | $l/2$ | 0 | $m/2$ | 0 |
| 3 | 0 | $m/2$ | 0 | $l/2$ |
| 4 | $M/2$ | 0 | $l/2$ | 0 |

where we assume that

$$M > l > m, \qquad \text{and} \qquad M + m > 2l, \tag{8}$$

so that an exchange between pairs of type 2 and 3 gives an outcome of $m$, one between 1 and 2 or 3 and 4 gives $l$, while one between 1 and 4 gives an outcome of $M$ and it is thus preferable. Moreover the matching $\{(1,4)(2,3)\}$ has a higher weight than the matching $\{(1,2)(3,4)\}$. Note that, given the structure of mutual compatibilities, 3-ways exchanges are not possible. In addition, even if we do not consider this possibility here, also 4-way exchange involving all the pairs has a lower reward than the two exchanges $(1,4)$ and $(2,3)$, due to the assumption in (8).

The state of the system at time $t$ is a subset of $\mathbb{N}^4$, and the simplicity of the chosen situation allows the explicit computation of the weight for the maximum weight matching once the composition of the pool is known. In fact, a maximum weight matching $m^*$ can be easily characterized. Since pairs of the same type are indistinguishable in this model, we define two matchings to be *equivalent*, if they prescribe the same number of exchanges between every fixed pair of types. The following theorem states that the maximum weight matching is essentially unique from this point of view.

**Theorem 2.** *Let $\boldsymbol{s} = (s_1, \ldots, s_4)$ be the state of the system, and let $G(\boldsymbol{s})$ be the corresponding graph. Then a maximum weight matching $m^*$ matches as much as possible pairs of types 1 and 4, then the remaining pairs (if any) of type 2 or 3 with the available pairs of type 1 or 4 respectively, and finally matches the remaining pairs of type 2 and 3. The corresponding maximum weight is therefore*

$$w(\boldsymbol{s}) = \begin{cases} M * s_4 + r \cdot \min\{s_1 - s_4, s_2\} + \\ \quad m * \min\{(s_2 - s_1 + s_4)_+, s_3\} & \text{if } s_1 \geq s_4; \\ M * s_1 + r * \min\{s_4 - s_1, s_3\} + \\ \quad m * \min\{(s_3 - s_4 + s_1)_+, s_2\} & \text{if } s_1 < s_4, \end{cases}$$

*where the notation $(x)_+$ stands for* $\max\{x, 0\}$.

*Proof.* Suppose that $a$ is a maximum weight matching not equivalent to $m^*$. We distinguish two cases: $s_1 \geq s_4$ and $s_1 < s_4$. We analyze the first case, the second one being analogous. As described above, $m^*$ matches the $s_4$ pairs of type 4 with the pairs of type 1. If the matching $a$ does the same, then it is easy to prove that $a$ is equivalent to $m^*$. Otherwise, suppose that there exists a pair $u$ of type 4 that is matched with a pair of type 1 through $m^*$, and is not matched with a pair of type 1 through $a$. Two possibilities occur: either the pair is left unmatched, or the pair is matched with a pair of type 3. Since we assumed $s_1 \geq s_4$, and that $a$ is a maximum weight matching, the first possibility leads to a contradiction and does not take place. This means that $u$ is matched with a pair of type 3, so that only two cases are possible: either an exchange of type $(1,4)$ having weight $M$ in $m^*$ is replaced by one of type $(3,4)$, having weight $l$, or two exchanges of type $(1,4)(2,3)$ in $m^*$ are replaced by two of type $(1,2)(3,4)$ in $a$. In both cases assumption (8) leads to a contradiction. This implies that $m$ is equivalent to $m^*$ in the case $s_1 \geq s_4$. □

We remark that in general an explicit formula for the maximum weight matching is not essential in the development of the simulations, since the maximum weight can be found using efficient algorithms [37]. In this case, we exploit the simple structure of the model in order to speed up the simulation.

### 5.1 Setting the parameters

The main difficulty in the study of the simulations is setting the parameters so to get a realistic model. We decided to fix one day as the time unit, tuning the other parameters consequently. Referring to the Italian situation, the following daily arrival rates are reasonable choices (see [51]):

$$\lambda_1 = 0.02, \quad \lambda_2 = 0.024, \quad \lambda_3 = 0.03, \quad \lambda_4 = 0.016.$$

Such rates are taken from [51], where the incompatible donor-recipient pairs participating to a kidney exchange problem are simulated on the basis of a suitable tree decision model. There are several factors influencing the distribution of blood types in the pool. In particular, as explained in [51], pairs in which the recipient is of type B and the donor is of type A occur more frequently than pairs in which the recipient is of type A and the donor is of type B, and the arrival rates in this paper have been chosen accordingly. We also assume that mature donors appear more frequently than young ones. The specific choice of the arrival rates gives rise to an unbalanced system, determining an accumulation effect. More specifically, in our simple model, the expected number of pairs of type 1 and 3 is higher than the expected number of pairs of type 2 and 4. Since pairs of type 1 and 3 can exchange only with pairs of type 2 and 4 and vice versa, this implies that pairs of type 1 and 3 tend to accumulate, and this property is confirmed in our simulations, as discussed below. Moreover, according to the chosen reward function (see the compatibility table), it is better to match type 1 with type 4 pairs, and type 2 with type 3 pairs, than to reverse these. As such, since the recipients of pairs of type 1, arriving at rate 0.02, do not exceed the total supply of $B$ donors (coming from pairs of type 2 and type 4 at rate 0.04), then all of type 1 will be matched. So, in the long run, only type 3 will accumulate.

More delicate issues are the choices of the discount rate, of the costs and of the compatibility values. One can justify the choice of the discount factor by explicitly considering the risk of death if the patient continues dialysis. We set $\delta = 0.9999$ (this choice coincides with the one made in [58]).

Concerning the costs, which are primarily intended to reflect monetary costs, we set the cost $c_s$ of a single exchange to be 1. This corresponds to $c_t = 0.5$ in (3), since $c_s = 2c_t$. Considering the fact that a kidney transplant costs approximately $100,000$ euros, $c_s$ corresponds to the monetary value $200,000$ euros. Therefore, setting the fixed costs $F = 0.1$, which we recall include all the expenses due to the maintenance and the management of the system for each round of exchanges, looks as a reasonable choice. On the other hand, the purely monetary costs of dialysis are about 50.000 euros per year. Hence, one day of dialysis is then about 140 euros, namely $c_u = 0.0007c_s$.

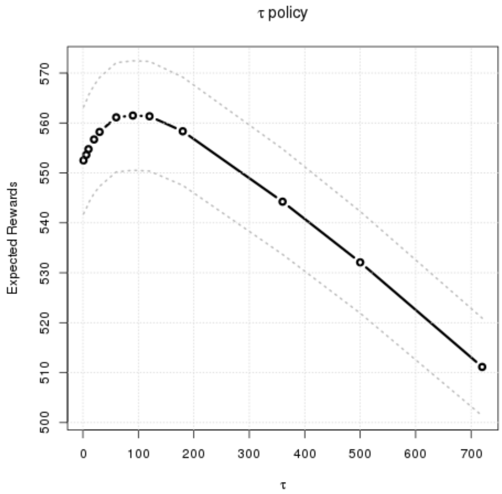
We are left to fix the immediate reward of an exchange of a given quality. The choice of this quantity is not obvious, and a possibility could be considering the Quality Adjusted Life Years (QALYs) expected by the patient as a result of the obtained kidney [58]. A meaningful numerical value for this quantity needs to reflect the cost of maintaining a patient on the waiting list for some time. We chose to set the reward of a single transplant to correspond approximately to the discounted cost of 30 years of dialysis, in the case the exchange is between pairs of type 1 and 4, and with 20 years for exchanges between pairs of types 1 and 2 and 3 and 4, and finally with 15 years of dialysis for exchanges involving pairs of type 2 and 3. Our final choices are therefore:

| $M$ | $l$ | $m$ |
|---|---|---|
| 9 | 7 | 6 |

### 5.2 Optimal strategy computation through Markov Decision Processes

In this section we compute the optimal performance derived through the solution of the MDP described in Section 3. The computational cost for generating and solving the MDP grows rapidly as a function of the problem size (number of patient types, waiting queue size, ...); hence the reported results have been derived by posing the following further constraints: four types of donor-recipient pairs, waiting queue size limited to 36 for each type, maximum number of new arrivals per time unit limited to 2. At any time, the MDP state represents the number of pairs in queue for each type, the possible actions in each state correspond to the choice of possible matchings (not necessarily maximal) that may be selected in such state. The state change after an action has been chosen is probabilistically defined on the basis of the arrival rates of each pair type, and on the queue size and maximum number of arrivals limits.

The number of MDP states for the given limits is $91,250$; it has been automatically derived from a high level model representation expressed using the Markov Decision Well-formed Net (MDWN) formalism [9, 10] to significantly reduce the probability of introducing errors, and to




| $\tau$ | Exp.Reward |
|---|---|
| 1 | $552.50 \pm 10.70$ |
| 6 | $553.62 \pm 10.76$ |
| 10 | $554.76 \pm 10.79$ |
| 20 | $556.70 \pm 10.83$ |
| 30 | $558.21 \pm 10.90$ |
| 60 | $561.13 \pm 10.94$ |
| 90 | $561.49 \pm 10.96$ |
| 120 | $561.35 \pm 10.98$ |
| 180 | $558.34 \pm 10.84$ |
| 360 | $544.24 \pm 10.56$ |
| 500 | $532.09 \pm 10.19$ |
| 720 | $511.14 \pm 9.78$ |

Figure 1: Results obtained in the simulations varying the time interval $\tau$, with waiting queue size limited to 36 for each type and maximum number of new arrivals per time unit limited to 2. The policies are evaluated through simulation in terms of the total expected $\delta = 0.9999$ discounted reward on a time horizon of ten years.

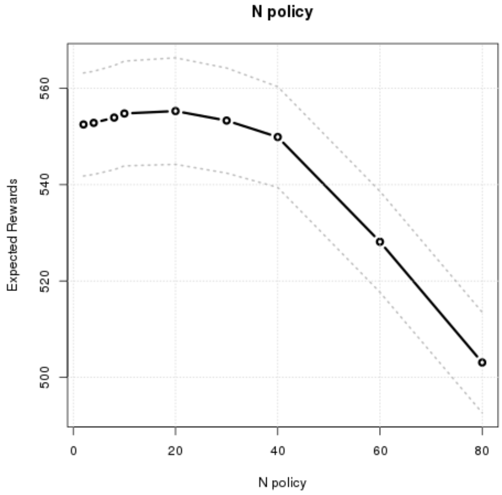
have a more parametric model from which different MDP could be derived for any given set of parameters.

The MDWN model comprises two subnets, the non deterministic one, generating all possible choices available in a given state, and the probabilistic one, generating the new arrivals as well as the discharge of donor-recipient pairs when the maximum queue length is reached. The MDP is generated by making the MDWN evolve (exhaustively) through the possible states, with an alternation of non deterministic *steps* and probabilistic *steps*. Intuitively, for each reachable state the set of all possible sets of pair associations compatible with that state (including the decision to postpone some association) define the possible actions and are built incrementally. For any given action, the next state is computed by selecting a few pairs to be dismissed (chosen randomly) if the queue size limit has been reached, and then by generating up to two arrivals, with probability distribution derived from the Poisson arrivals of each pair type and taking into account the ratio between the arrival rate of each type w.r.t. the total arrival rate.

MDP generation from MDWN model requires, on an Intel i7 processor, less than 1 hour and 7.1Gb of memory; while its solution through policy iteration more than 1 day and 12.6Gb.

With the same constraints, we estimated the expected reward implementing policies $\pi_\tau$ and $\xi_N$ for different values of $\tau$ and $N$ on simulated sequences of arrivals over a time horizon of ten years. The simulated experiments were executed using an ad-hoc developed simulator whose functionalities will be described in the next section. Confidence intervals were computed for each selected index, setting the confidence level to 95%. The expected reward obtained in these simulations is plotted in Figs. 1 and 2.

The expected reward obtained by MDP solution (i.e. 563.68) is higher than the expected rewards computed through simulation, however it falls within the confidence interval derived through simulation for several values of $\tau$ and $N$. Moreover, it is important to highlight that the computation of the performance of almost myopic policies through simulation requires substantially less computational resources (i.e. execution time: 5s. and memory used: 4Mb), so that it can be applied to more complex cases, e.g. when the constraint on the queue size is removed (see next section) and the solution of the MDP is not possible anymore. Finally, we




| $N$ | Exp.Reward |
|---|---|
| 2 | $552.50 \pm 10.70$ |
| 4 | $552.83 \pm 10.71$ |
| 8 | $553.93 \pm 10.81$ |
| 10 | $554.78 \pm 10.88$ |
| 20 | $555.28 \pm 11.06$ |
| 30 | $553.31 \pm 10.91$ |
| 40 | $549.89 \pm 10.46$ |
| 60 | $528.14 \pm 10.48$ |
| 80 | $503.07 \pm 10.44$ |

Figure 2: Results obtained in the simulations varying the pool size required for matching with waiting queue size limited to 36 for each type and maximum number of new arrivals per time unit limited to 2. The policies are evaluated in terms of the total expected $\delta = 0.9999$ discounted reward on a time horizon of ten years.

observed that the results of almost myopic policies obtained removing the constraints on the waiting list size are similar to those in Figs. 1 and 2.

# 6 Kidney exchange simulator

In this section we present a new simulation tool that we developed to study the performance of almost myopic policies on more realistic kidney exchange models.

The tool is implemented in C++ programming language and it is structured around two main classes, i.e. the class implementing the core functionalities of the simulator and the class implementing the kidney exchange model. This modularization in the simulator code facilitates the integration of new simulator functionalities or model features. Moreover, the simulation tool exploits the multi-platform LpSolve library (*http://lpsolve.sourceforge.net*) to efficiently solve the optimization problem arising at each exchange round.

The simulator input parameters are divided in two sets: those related to the simulation itself and those related to the description of the kidney exchange model. In details, for the first set the user has to specify on the command line the following parameters: a)the simulation duration, b)the transient phase in which simulation statistics are not gathered, c)the total number of simulation runs, d)the confidence level to be used in the parameters estimation and e)the simulation seed.

The second set of parameters describes the kidney exchange model in terms of its costs (i.e. Fixed Cost, Patient Cost, Transplant cost) and rewards, number of time steps between two batch arrivals[1], distribution of the batch size at each arrival, distribution of types of arrived pairs, pair compatibility degree, and the type of almost myopic policy under investigation with the required parameters. These parameters are stored in a file that is passed as input to the simulator. Each model parameter in the file is identified by a specific tag whose name was chosen to be self explanatory. Taking the model description as an input parameter, rather than embedding it into the code, allows modellers to study the almost myopic policies on different scenarios. We

[1]Time is discretized and the arrival of groups of up to k pairs are considered every n time steps.

```
#INPUT FILE#
Discount : 0.9999

#Total number of pairs#
Pairs : 4
#Arrival Frequency#
Frequency : 1

#Distribution of new pairs for each arrival:#
# 1 pair with 0.082253807#
# 2 pair with 0.003815008#
# no pair with 0.913931185#
Arrivals: 0.082253807 0.003815008

#Distribution of pair types#
Probabilities: 0.2222 0.2667 0.3333 0.1778

FixedCost : 0.1
PatientCost : 0.0007
TransplantCost : 0.5

#Total number of exchanges#
Exchanges : 4

Rewards : 7 6 9 7

#Exchange compatibility in list form#
Constraints :
1: 1 3
2: 1 2
3: 2 4
4: 3 4
```

```
#INPUT FILE#
Discount : 0.9999

#Total number of pairs#
Pairs : 4
#Arrival Frequency#
Frequency : 1

#Distribution of new pairs for each arrival:#
# 1 pair with 0.082253807#
# 2 pair with 0.003815008#
# no pair with 0.913931185#
Arrivals: 0.082253807 0.003815008

#Distribution of pair types#
Probabilities: 0.2222 0.2667 0.3333 0.1778

FixedCost : 0.1
PatientCost : 0.0007
TransplantCost : 0.5

#Total number of exchanges#
Exchanges : 4

#Exchange compatibility in matrix form#
Matrix:
0 0 0 0
7 0 0 0
0 6 0 0
9 0 7 0
```

Figure 3: Model description files for the simplified model in Section 5

next briefly discuss how to deal with the compatibility constraints. There are different ways to provide the pairs compatibility information as input to the simulator: they require to specify all the admissible exchanges. More precisely, let $K$ be the maximum length of the admissible cycles (in our experiments $K = 2$ or 3): starting from the compatibility matrix $R$ defined in Section 2, the information provided to the simulator can be formalised as a set $\{e_1, \dots, e_{ne}\}$ of all legal exchanges, where each $e_i$ is a sequence of pair types of length $k$, with $k \in \{2, \dots, K\}$, which represents an admissible $k$-way exchange. More precisely, for each $e \in \{e_1, \dots, e_{ne}\}$, there exists $k \in \{2, \dots, K\}$ and $t_0, \dots, t_{k-1}$ in $\mathcal{T}$ such that

$$e = (t_0, \dots, t_{k-1}) \quad \text{and} \quad R_{t_l, t_{(l+1) \bmod k}} > 0 \text{ for every } l \in \{0, \dots, k-1\}. \tag{9}$$

Note that an equivalence relation exists among sequences of pair types: two sequences are equivalent under rotation, meaning that if one exchange $e$ can be obtained from $e'$ by applying a rotation to its elements (e.g. $e = (t_0, \dots, t_{k-1})$, $e' = (t_2, ..., t_{k-1}, t_0, t_1)$), they describe the same $k$-way exchange. The reward associated with an exchange $e = (t_0, \dots, t_{k-1})$ is equal to the sum of the rewards of each transplantation in the chain:

$$r(e) = \sum_{l=0}^{k-1} R_{t_l, t_{(l+1) \bmod k}} \tag{10}$$

The first format for specifying the possible exchanges in the simulator input file requires to provide the number of exchanges $ne$, their reward, and for each pair type the list of exchanges where it is involved. The input file corresponding to the simplified model introduced in Section 5 according to this format is shown in Fig. 3(left): the number of exchanges that can be considered

Table 1: Distribution of pair types.

| Blood Type | Recipient 0 | | Recipient A | | Recipient B | | Recipient AB | |
|---|---|---|---|---|---|---|---|---|
| Donor 0 | [1] | 0.14 | [2] | 0.063 | [3] | 0.024 | [4] | 0.005 |
| Donor A | [5] | 0.378 | [6] | 0.068 | [7] | 0.061 | [8] | 0.005 |
| Donor B | [9] | 0.12 | [10] | 0.051 | [11] | 0.012 | [12] | 0.002 |
| Donor AB | [13] | 0.02 | [14] | 0.028 | [15] | 0.021 | [16] | 0.001 |

follows tag *Exchanges*, the list of rewards associated with each exchange follows tag *Rewards*), while tag *Constraints* is followed by $n$ lists, one for each pair type, indicating the (multi)set of exchanges in which that pair can be involved (it is a multiset because two or three pairs of the same type may be involved in an exchange, assuming the possibility that other factors not explicitly represented in the model may justify the need for such an exchange).

An alternative representation for the input file when only 2-way exchanges are admitted (i.e. $K = 2$), is a lower triangular matrix $R'$ defined as follows

$$R'_{i,j} = \begin{cases} R_{i,j} + R_{j,i} & \text{if } i \geq j \text{ and } R_{i,j}R_{j,i} > 0 \\ 0 & \text{otherwise.} \end{cases}$$

For instance, the file corresponding to the simplified model introduced in Section 5 is shown in Fig 3(right) where after indicating the number $m$ of possible exchanges (tag *Exchanges*) the compatibility is specified through a matrix of size $n \times n$ (tag *Matrix*). The number of non-null elements in the matrix matches the number of possible exchanges. The position of these elements indicates the pair types involved in the exchange, while the value indicates the reward associated with that exchange.

**Specifying the pair types.** The set of possible pair types can be defined taking into account several factors, besides blood type, that may influence the compatibility or the quality of the exchange (see e.g. [25, 31, 38, 32, 41, 42, 39] for studies on this issue). The experiments in next section for instance focus on the age of donor and recipient: this introduces a finer classification in types induced by a distinction of the donors and the recipient into two classes: *young* and *mature* (see also [40]). This is a relevant issue, since it has been observed that exchanges involving young donors [43], or with a good match between donor and recipient age ensure an improved graft survival [36].

As discussed in next section in our experiments we consider three scenarios with the following characterization of pair types:

1. based only on donor and recipient blood types (which leads to 16 pair types);
2. based on donor and recipient blood type and donor age (which leads to 32 pair types);
3. based on donor and recipient blood type and donor and recipient age (which leads to 64 pair types).

In all considered scenarios, according to the assumptions made in the description of the model, the pairs of type $i$ arrive according to a Poisson process with rate $\lambda_i$ and join the exchange system. Actually the simulation time is discretized at steps of length $\Delta t$ (in our experiments $\Delta t$ = one day), up to $k$ arrivals are generated once every $n$ time steps, and the probability of having $k$ arrivals (up to a predefined maximum) is computed according to a Poisson distribution.

Table 2: Quality of transplants based on donor's age: $M > m$

| Donor's age | Quality |
|---|---|
| Young Donor | M |
| Mature Donor | m |

Table 3: Quality of transplants based on donor's and recipient's age: $M > l > m$

| Age | Young Recipient | Mature Recipient |
|---|---|---|
| Young Donor | M | l |
| Mature Donor | m | M |

There are several factors influencing the distribution of blood types in the pool. Referring to the Italian situation, the daily arrival rates in Table 1 are reasonable choices for blood type pairs. Such rates are taken from [51] (see also [26, 52, 57]) where the incompatible donor-recipient pairs involved in a kidney exchange problem are simulated on the basis of a suitable decision tree model. Moreover, considering the age aspect, the subclasses arrivals are derived by considering that mature donors and mature recipient are more frequent than young ones, and in our experiments they are the 63.2% of the population.

The specific choice of the arrival rates may lead to an unbalanced system, determining an accumulation effect of some pair types. This was already observed in the small example in the previous sections, and will be discussed with more details in the comments on our experiments.

**Specifying compatibility and quality of exchanges.** The definition of the pairs compatibility and exchanges quality as input is another important aspect of the simulator that provides a tool to evaluate the impact of the exchange quality chosen value on the pairs selection at each decision point. For instance in the three models studied in the next section, the quality of a $k$-way exchange, for $k \in \{2, 3\}$, is given by the sum of the quality of each transplant in the exchange set. In details in the first scenario considered in next section we assume that all the transplants have the same quality index. In the second scenario, the quality of a transplant depends on the age of the donor, see Table 2. In the the third scenario, the quality of a transplant depends both on the age of the donor and of the recipient, according to Table 3. In all scenarios, blood types do not influence the quality of transplants, but only determine the compatibility.

**Integer Linear Programming model for computing maximal matching** As mentioned before, at each exchange round the simulator solves an optimization problem to compute the maximal matching. There are several solutions available in the literature, see e.g. [1] and references therein. Here we introduce an integer linear programming problem, whose dimension is not affected by the number of pairs in the system, but only by the number of types and the compatibility structure. In the simulations, the resulting integer programming problem will be solved using the LpSolve library.

Let us consider again the set $\{e_1, \ldots, e_{ne}\}$ of possible exchanges introduced earlier: each $e_i$ represents a possible $|e_i|$-way exchange. Let $\boldsymbol{s} = (s_1, \ldots, s_n)$ be a state of the system, indicating how many pairs of each type are currently present in the system. An optimal matching can be found by solving the following integer linear programming problem, with $ne$ integer variables $x_1, \ldots, x_{ne}$, indicating how many exchanges of type $e_1, e_2, \ldots e_{ne}$ should be performed in state

$\boldsymbol{s}$. The objective function (to be maximized) is:

$$f(x_1, \ldots, x_{ne}) = \sum_{j=1}^{ne} r(e_j) x_j$$

under the following constraints:

$$\forall i \in \{1, .., n\}, \quad \sum_{j=1}^{ne} x_j \ \mathrm{count}(e_j, i) \leq s_i$$

where $r(e_j) = \sum_{h=0}^{k-1} R_{t_h, t_{(h+1) \mathrm{mod} k}}$, with $k = |e_j|$, is the reward associated with the $k$-way exchange $e_j$ and function $\mathrm{count}(e_j, i)$ denotes the number of occurrences of type $i$ in $e_j$. As mentioned at the beginning, the number of variables of the optimization problem only depends on the number of admissible exchanges, while the number of pairs in the system affects the optimization problem only changing the constants in the constraints.

**Fairness selection when several optimal solutions exist.** Since several optimal solutions can exist for an exchange round in given system state, one needs to define an appropriate method to select one of them. In particular, in our tool we refined the objective function of the optimization problem with the aim of balancing the waiting times of all pair types and mitigate starvation of under-demanded pairs. Hence in the optimization problem the objective function has been modified so that a slightly higher reward is associated to the exchanges involving pair types that have experienced longer waiting times. This variation is applied only when defining the optimization problem to be solved by *LpSolve*, not in the computation of the global reward index estimated by the simulator. Note that this modification allows to select one among the maximal matchings, and does not introduce new solutions. More specifically, the user-defined reward associated with each exchange is thus incremented by a suitable function of the waiting times of each involved type, for instance $R^w_{ij} = R'_{ij} + \mathcal{C} * WT_i + \mathcal{C} * WT_j$ where $R'$ is the compatibility matrix derived from the input file, and $WT_i$ is the median of the waiting times of type $i$. This ensures that types with a longer waiting time median will be favoured; it is important to choose a small enough $\mathcal{C}$ constant, so that the selected optimal solution surely belongs to the set of optimal solutions computed without increments.

We now consider the modified integer linear programming problem and show that it is possible to introduce a perturbation based on observed waiting times without changing the nature of the achieved optimal solution. In the following discussion we shall refer again to the graph-based definition of the matching problem for simplicity in the proof, but it is not difficult to map the graph based model into the corresponding ILP model. We recall that a maximum weight matching is a set of cycles in a directed graph, that do not share nodes, and such that no other set of disjoint cycles exists with greater overall weight (see Definition 1).

**Lemma 1.** *Let $V$ be a finite set, let $E \subset V^2$, and let $v : E \to [0, +\infty[$. Let $G = (V, E, v)$ be the corresponding directed weighted graph. Let $\epsilon : E \to [0, +\infty[$, and let $G_\epsilon = (V, E, v + \epsilon)$. Let $\bar{w}$ be the weight of a maximum weight matching (possibly with constrained maximum cycle length) in $G$, and let $\underline{O}$ be the set of not optimal matchings in $G$. Let us define $d$ as the minimum difference between the weight of any maximal weight matching and the weight of any other non-optimal matching:*

$$d = \bar{w} - \max_{m' \in \underline{O}} \sum_{(i,j) \in m'} v_{ij} \tag{11}$$

*where $m'$ is a set of arcs in $G$ defining a matching (i.e. the union of the arcs belonging to the k-way disjoint cycles composing matching $m'$). Then, if* $\max_{(i,j)\in E}\epsilon_{ij} < d/|V|$, *any maximum weight matching in $G_\epsilon$ is a maximum weight matching in $G$.*

*Proof.* First note that the weight of a maximum weight matching in $G_\epsilon$ is greater than or equal to $\bar{w}$ since $\epsilon$ is positive. Let $m' \in \underline{O}$. The weight of the matching $m'$ in $G_\epsilon$ is $\sum_{(i,j)\in m'}(v_{ij}+\epsilon_{ij})$. Note that $|m'| \leq |V|$ since at most we can have one arc departing from each node involved in a cycle composing $m'$, hence $\sum_{(i,j)\in m'}\epsilon_{ij} < |m'|d/|V| \leq d$. Therefore the weight of $m'$ in $G_\epsilon$ satisfies

$$\sum_{(i,j)\in m'}(v_{ij}+\epsilon_{ij}) < \sum_{(i,j)\in m'} v_{ij} + d \leq \bar{w},$$

and thus $m'$ is not optimal in $G_\epsilon$. □

In our experiments, we use the trick described in the previous lemma in order to favour solutions which exhibit fairness among the optimal ones, and, for compatible pairs $i$ and $j$ with $i$ and $j$ in $\{1,\ldots,n\}$ we set

$$\epsilon_{ij} = c * (W_i + W_j)/N,$$

where $N$ is (an upper bound for) the total number of pairs in the waiting list, $W_i$ and $W_j$ are (upper bounds for) the median waiting times of pairs of type $i$ and $j$ and $c$ is a small enough constant chosen in order to verify the assumptions of Lemma 1,

A possible choice for $c$ is $c = \delta/2m\mathcal{W}_u\mathcal{E}_u$ where $\delta < d$, with $d$ defined as in (11), $\mathcal{W}_u$ is an upper bound for the waiting time and $\mathcal{E}_u$ is an upper bound on the number of possible exchanges in the considered state, then the conditions imposed by the lemma are surely satisfied. If the rewards in the original optimization problem are all integer[2] we can safely choose any $\delta < 1$ (e.g. $\delta = 0.9$).

**Dealing with the accumulation of under-demanded types in $\xi_N$ policy.** The accumulation of so-called under-demanded pairs was observed in [54] and discussed in Section 4; when a given pair type starts accumulating it may force the quasi myopic policy $\xi_N$ to organize an exchange round at each time step (so that it becomes a myopic policy). To avoid this effect in our simulation tool we have explicitly implemented the possibility of specifying a set of pair types (associated with tag *NoAccumulation* in the input file) whose number is not considered in choosing when an exchange should be scheduled in policy $\xi_N$: this allows to define set $\mathcal{O}$ of Eq.7.

**Computing measures of interest.** The simulator output is saved in two files: in the first file, with extension .output, a set of measures of interest for each run is reported in textual form. In the second file (with extension .cvs) the same set of measures associated with their confidence intervals is saved as comma-separated values. In details for each run the following measures are computed: a)average exchange reward for time unit; b)average exchange reward for matching; c) average number of discharged patients per time unit; c) average number of discharged patients per exchange round; d) average number of pairs in hospital per each type; e) average waiting time per each type; d) ratio between the total number of discharged patients and the total number of admitted patients; e) expected discounted reward; d) the number of pairs in the hospital of each type at the end of simulation.

[2]The integer rewards hypothesis is not restrictive: as a matter of fact the only requirement is to have finite precision rewards, which could be transformed into integers multiplying them by a suitable factor.

The simulator is available free of charge for academic institutions and non profit organizations; it can run on Linux, Unix and MacOS X system: `http://www.di.unito.it/~beccuti/software.html`

# 7 Evaluation of almost myopic policies on more complex scenarios

In this section we exploit our parametric simulator to evaluate the performance of the two introduced classes of almost myopic policies in the three scenarios outlined in the previous section. The main aim is therefore to derive some insight on how to choose the best timing between two exchange rounds, or, for state dependent policies, a suitable threshold on the optimal number of pairs in the waiting list (possibly filtering out accumulating pair types) that should be used to trigger a matching. Moreover, we performed an experimental sensitivity analysis to enlighten the role and the effects of the various parameters involved. Indeed, the choices of the discount rate, of the costs, and of the compatibility values are a delicate issue.

In details, in our experiments we assume the arrival rates as described in the previous section, 1 day as time unit, $3,600$ days as the simulation horizon, and 0.9999 as discount factor. Concerning the model costs, which are primarily intended to reflect monetary costs, we will report several experiments for different choices of the related parameters. In particular, according to the consideration reported in Subsection 5.1, we assume as baseline/reference costs the following values $c_t = 0.5$, $F = 0.1$, and $c_u = 0.0007$, and then we introduce small variations of these values to study their effects on the simulation results.

With considerations similar to those of Subsection 5.1, we fixed the following choices for the immediate reward of a transplant, in the three considered models (introduced in Section 6):

**Model 1:** in this case all the transplants are of the same quality. We fixed the immediate reward of each transplant equal to 4.5 (hence an exchange involving two pairs has reward 9) corresponding approximately to the discounted cost of 30 years of dialysis;

**Model 2:** in this case the quality of the transplant depends on the age of the donor (see Table 2). We fix $M = 4.5$, which again corresponds approximately to the discounted cost of 30 years of dialysis, and $m = 3.5$ which corresponds to 20 years;

**Model 3:** in this case the quality of the transplant depends on the age of the donor and of the recipient (see Table 3). Reasoning as in the previous cases, we fix $M = 5$, $l = 4$, and $m = 3$.

The performance of the policies are measured and compared in terms of the total expected discounted reward over $1,000$ runs for each experiment, and confidence intervals are computed setting the confidence level to 99%. We ran the simulations for the three models, with the following choice of parameters:

1. time dependent almost myopic policies for $\tau \in \{15, 30, 60, 90, 120, 150, 180, 225, 240, 300, 360, 400, 450, 600, 720\}$;

2. state dependent almost myopic policies for $N \in \{2, 4, 8, 10, 12, 14, 20, 22, 24, 30, 32, 34, 40, 60, 80\}$;

Finally, Table 4 summarized how the performed experiments are organized with respect to the three models and all parameters.

| | **Model 1** | **Model 2** | **Model 3** |
|---|---|---|---|
| Policy $\pi_\tau$ | X | X | X |
| Policy $\xi_N$ | X | X | X |
| Varing $F$ cost (i.e. 0.1, 0.5 and 1) | X | X | X |
| Varing $C_s$ cost (i.e. 0.1 and 1) | X | X | X |
| Varing $C_u$ cost (i.e. 0.00007,0.0014,0.0007). | X | X | X |
| 2-way exchanges | X | X | X |
| 2-way and 3-way exchanges | X | | |
| with accumulated pairs for $\xi_N$ policies | X | X | X |
| without accumulated pairs for $\xi_N$ policies | X | | |
| without final match for $\xi_N$ policies | X | X | X |
| with final match for $\xi_N$ policies | X | | |

Table 4: Summary of the 1,048 experiments performed.

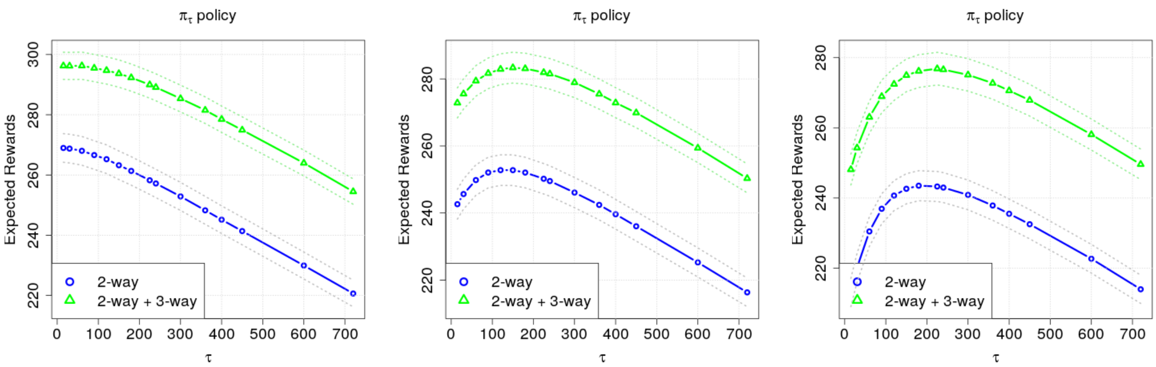



Figure 4: $\pi_\tau$ policy comparison for Model 1 with $c_s = 1$ $c_u = 0.0007$, and varying $F$ (i.e. 0.1, 0.5 and 1).

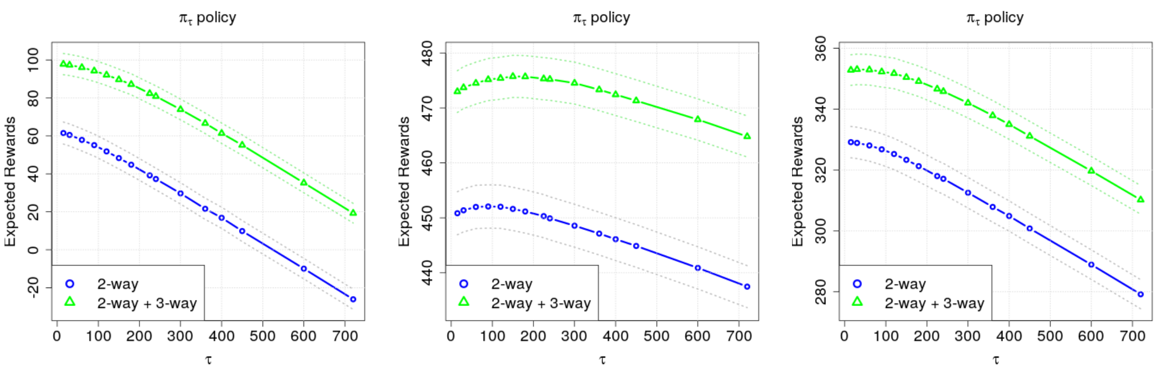



Figure 5: $\pi_\tau$ policy comparison for Model 1 with: $c_s = 1$, $F = 0.1$, and $c_u = 0.0014$(left), $c_s = 1$, $F = 0.1$, and $c_u = 0.00007$ (center), and $c_s = 0.1$, $F = 0.1$, and $c_u = 0.0007$(right).

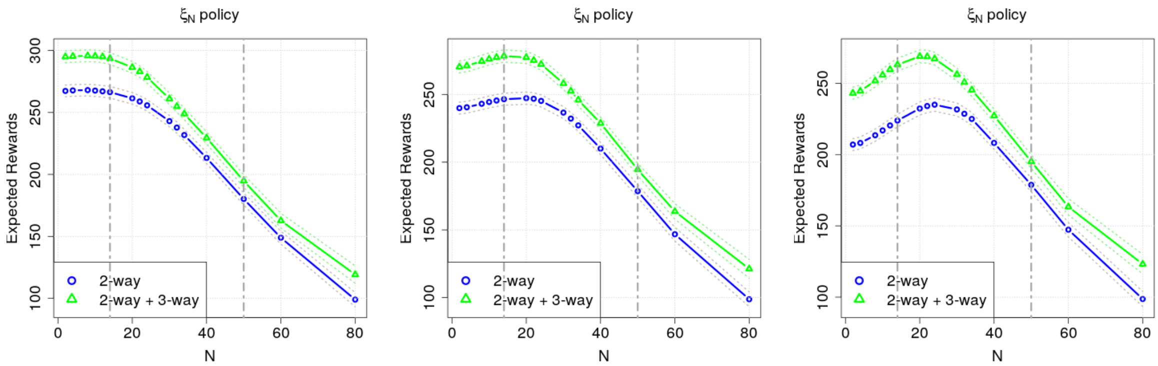



Figure 6: $\xi_N$ policy comparison for Model 1 with $c_s = 1$ $c_u = 0.0007$, and varying $F$ (i.e. 0.1, 0.5 and 1).

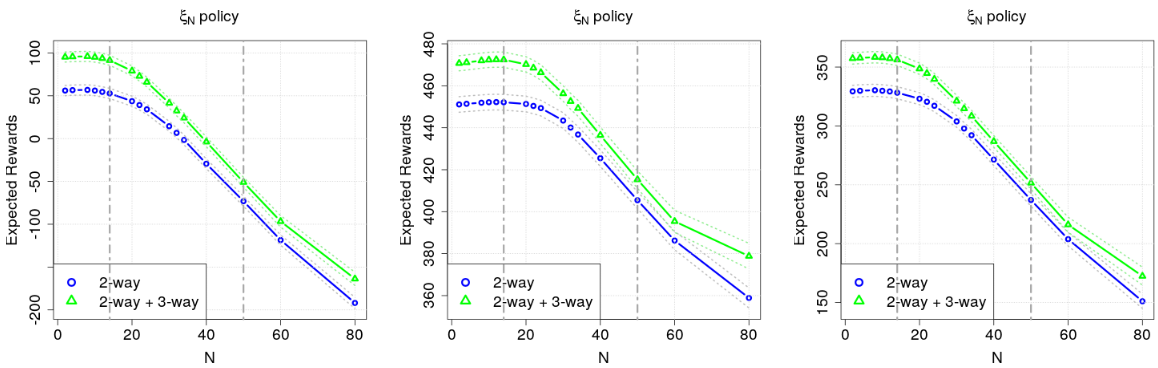



Figure 7: $\xi_N$ policy comparison for Model 1 with : $c_s = 1$, $F = 0.1$, and $c_u = 0.0014$(left), $c_s = 1$, $F = 0.1$, and $c_u = 0.00007$ (center), and $c_s = 0.1$, $F = 0.1$, and $c_u = 0.0007$(right).

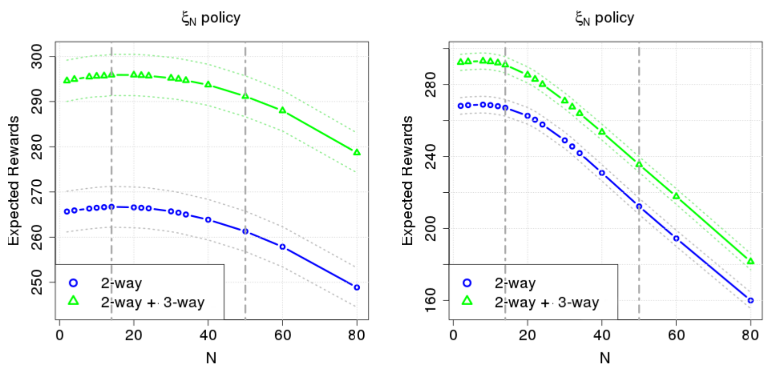



Figure 8: $\xi_N$ policy for Model 1 with $c_s = 1$, $F = 0.1$, and $c_u = 0.0007$ considering accumulation of under-demanded type(left), and $\xi_N$ policy for Model 1 with $c_s = 1$, $F = 0.1$, and $c_u = 0.0007$ assuming a final match before ending the simulation(right).

**Policy comparison for Model 1** We start our discussion considering the results obtained by the first model (i.e. based only on donor and recipient blood types) taking into account 2 and 3-way exchanges. We also recall that in this model all the transplants have the same quality (i.e. reward). In Fig. 4(left) the results of our simulations with respect to the baseline costs and $\tau$ policy are shown. With this choice of parameters, the total discounted expected reward is approximately a decreasing function of $\tau$, for the policies performing only 2-way exchanges. This suggests that, in this setting, it is better to execute exchanges as soon as they become available. This behavior is less emphasized for the policies allowing 3-way exchanges.

The other two graphs in Fig. 4 show that if the fixed cost $F$ is increased (i.e. 0.25 and 1 resp.) then the choice of waiting between a matching and the subsequent one looks more important because the maintenance and the management of the system makes it more convenient to wait for more than one transplant for each round of exchanges. For instance for $F = 0.5$, i.e. Fig. 4(center), the reward is maximized for $\tau$ close to 90, while for $F = 0.1$ the reward is maximized for $\tau$ close to 225.

In Figs. 5(left) and 5(center) the results of our simulations varying the costs of dialysis $c_u$ are reported. As effect of decreasing $c_u$ we obtain a better reward increasing the waiting time between one exchange and the other. However this effect is less evident than one obtained by increasing $F$ because all the transplants have the same quality.

The variation of the cost $c_s$ of a single exchange does not affect the results trend: the reward values are all scaled proportionally. In Figs. 5(right) the results obtained assuming $c_s = 0.1$ are shown, while all the other costs are fixed to the baseline values.

In Fig. 6(left) the results of our simulations with respect to the baseline costs and $N$ policy are shown. In this situation the policies $\xi_N$ performing only 2-way exchanges, as well as the ones allowing 3-way exchanges, show that better rewards are obtained for $N$ values between 2 and 14. Fig. 6 suggests that when the Fixed cost $F$ is increased to 0.25 the reward is slightly better for $N$ values between 10 and 14. Instead fixing $F = 1$ a non monotonic trend appears as shown for $\tau$ policy, and better rewards are derived for $N$ close to 24. Moreover if we focus on the N values between the two dashed vertical lines in Figs. 6, in which the average waiting time between two matches is in the same range of the $\tau$ values considered in $\pi_\tau$ experiments, conclusions similar to that proposed for $\tau$ policies can be derived.

The variation of the $c_u$ cost or $c_s$ cost does not affect the general shape of the resulting curve as shown in Fig 7, but the obtained reward values are substantially different.

The plot in Fig. 8(left) shows how $\xi_N$ policy rewards for the baseline case are affected by including under-demanded types (i.e. donor A, recipient O and donor A, recipient O) for the count triggering the exchanges. As expected, the match frequency is substantially increased due to an accumulation effect making less significant the reward difference when varying $N$. For instance the average waiting time between two matches for $N = 50$ is $\sim 50$ times smaller than the same measure computed excluding under-demanded types.

Fig. 8(right) reports instead the results obtained by the baseline case introducing a final matching before ending the simulation. This makes the rewards a bit higher, in particular for higher values of N. Indeed, increasing $N$ we have an higher probability of having possible matches at the simulation end.

Finally, comparing the maximum total discounted expected reward, we observe that, as expected, both policies allowing for 3-way exchanges give better results with respect to those involving only 2-way exchanges. Moreover a higher maximum total discounted expected reward is obtained by $\tau$ policies than $\xi_N$ policies this can be partially justified by the effect of under-demanded pairs.

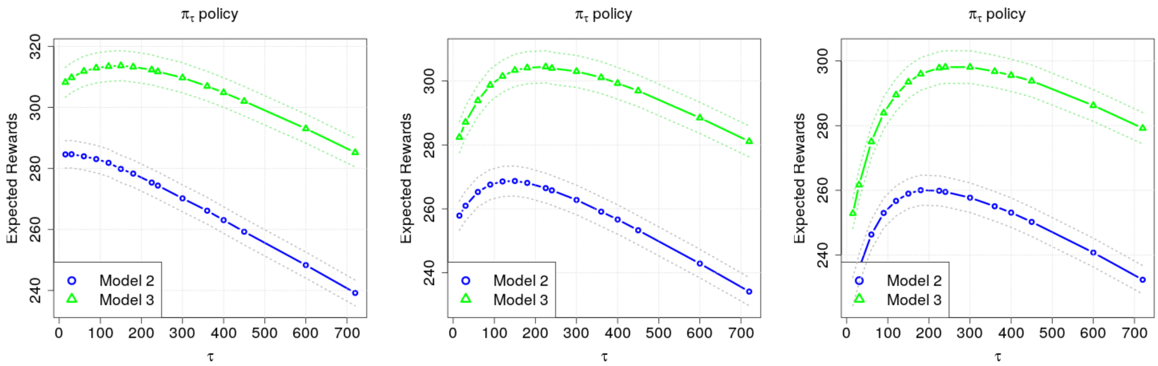



Figure 9: $\pi_\tau$ policy comparison for Model 2 and 3 with $c_s = 1$ $c_u = 0.0007$, and varying $F$ (i.e. 0.1, 0.5 and 1).

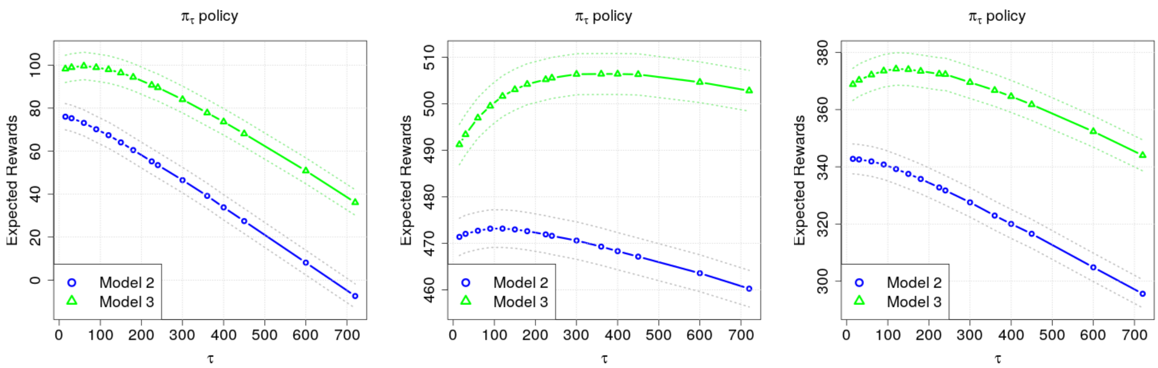



Figure 10: $\pi_\tau$ policy comparison for Model 2 and 3 with: $c_s = 1$, $F = 0.1$, and $c_u = 0.0014$(left), $c_s = 1$, $F = 0.1$, and $c_u = 0.00007$ (center), and $c_s = 0.1$, $F = 0.1$, and $c_u = 0.0007$(right).

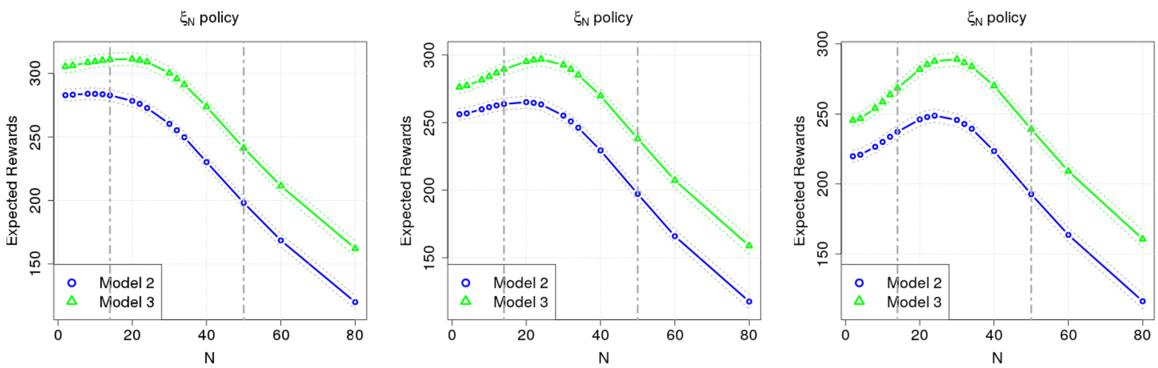



Figure 11: $\xi_N$ policy comparison for Model 2 and 3 with $c_s = 1$ $c_u = 0.0007$, and varying $F$ (i.e. 0.1, 0.5 and 1).

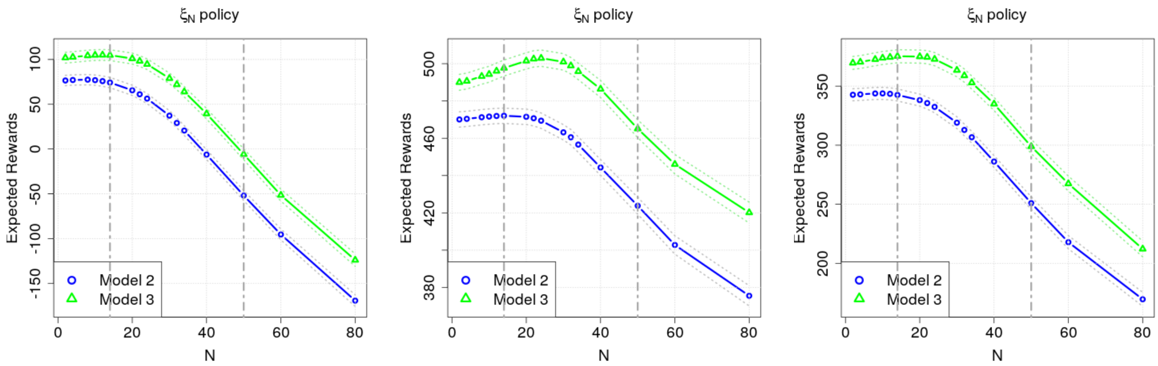



Figure 12: $\xi_N$ policy comparison for Model 1 with : $c_s = 1$, $F = 0.1$, and $c_u = 0.0014$(left), $c_s = 1$, $F = 0.1$, and $c_u = 0.00007$ (center), and $c_s = 0.1$, $F = 0.1$, and $c_u = 0.0007$(right).

**Policies comparison for Model 2 and 3** In this kidney exchange models differently by the Model 1 we take into account the quality of the executed transplants assuming different rewards (see previous section for detailed reward definition). While overall the comments are similar to those made for Model 1, one characteristic which distinguishes the results of the simulations for Model 2 and 3 from the ones for Model 1, is the fact that in general, in this case, it seems better to wait to perform transplants of higher quality.

As before, we start analyzing our default setting for $\tau$ policies. The total discounted expected reward is again a decreasing function of $\tau$ for Model 2, while for Model 3, where more pair types are involved, the benefits that can derive from waiting are more evident. In this case, as shown in Fig. 9(left) the total discounted expected reward is increasing for $\tau$ varying from 120 to 150, and then it starts to decrease.

As reported by Figs. 9(center) and 9(right) this latter tendency is observed also for Model 2 increasing the Fixed cost $F$ to 0.5 and 1. However, the benefits of waiting are still more amplified for Model 3, that shows a greater variation among the rewards. For instance assuming $F = 1$, better total discounted expected rewards are obtained for $\tau$ varying from 225 to 300 for both models.

In Figs. 10(left) and 10(center) the results of our simulations varying the costs of dialysis $c_u$ are reported. Differently by Model 1 the $c_u$ variation has a stronger effect on the rewards. Indeed the advantages of waiting in Model 3 are lost if $F$ is 0.0014 (see Figs. 10(left), while they become much more evident, in both the models, when $F$ is equal to 0.00007.

Moreover, as in the Model 1, the variation of the cost $c_s$ of a single exchange does not affect the result trend. Indeed, Fig. 10(right) shows that all the reward values are scaled proportionally with respect to those reported in Fig. 9(left) for the baseline case.

For $\xi_N$ policies the simulation results, shown in Figs 11 and 12 , come to the same conclusions already discussed for $\tau$ policies. This is more evident if we focus on the N values between the two dashed vertical lines in which average waiting time between two matches is in the same range of the considered $\tau$ values.

In both policies we observe that better results are obtained when the exchange quality takes into account the blood type and the age of the the donor and recipient (i.e. Model 3). This can be justified by the fact that in this case exchanges with higher rewards can be favoured.

Moreover a higher maximum total discounted expected reward is still achieved by $\tau$ policies than $\xi_N$ policies, even if this trend is less evident with respect to the Model 1.

## 8 Conclusions

In this paper we showed how the *dynamic* kidney exchange problem can be conveniently modelled by a MDP model taking into account the quality of the performed exchanges as well as the organizational cost. However, the complexity of the MDP, in terms of possible states and available decisions, prevents the possibility to compute the optimal policy in real cases, so that only heuristic approaches can be exploited to obtain a sub-optimal solution. Among these heuristic approaches we focused on *almost myopic* policies, because they are currently adopted by many transplant organizations.

In particular, a time dependent and a state dependent almost myopic policy were proposed and studied in this paper. The performance of these two almost myopic policies were first estimated through simulation and compared with the optimal policy obtained by numerical solution of the MDP model in a simplified case. The obtained results show that the performance of these two policies is quite similar to that of the MDP optimal policy.

Then, these two policies were investigated in more realistic scenarios thanks to a parametric simulator (freely available for download at `http://www.di.unito.it/~beccuti/software.html`) which allows users to consider multiple characteristics of both donors and recipients, different degrees of compatibility (quantifying the quality of each possible exchange) and different cost factors concurring to the definition of the reward function on which the evaluation is based. This simulation model and parameters are described in the paper, showing how the simulator can be used to study different scenarios that can be configured by modifying the rich set of input parameters. The results of several experiments are proposed to show how (1) different scenarios can be configured and (2) varying the parameter values it is possible to investigate their impact on the overall performance (e.g. the kidney transplant costs, dialysis costs, ...) i.e. on the achievable reward. The study of other sub-policies as well as their integration in our simulator will be carried out as future works. Moreover, models taking into account other aspects such as tissues compatibility and/or body mass index will be investigated.

## Acknowledgments

The authors wish to thank Fioravante Patrone, Vito Fragnelli and Lina Mallozzi for helpful suggestions and discussions, and Nicola Rebagliati for his contribution to the numerical simulations.

The work of Silvia Villa has been supported by the Regione Basilicata project "Analisi strategica, modellizzazione matematica ed algoritmi per un programma di scambi di organi" and it has been carried out in collaboration with "Centro Nazionale Trapianti". The research work of Giuliana Franceschinis has financial support of the Università del Piemonte Orientale.

## References

[1] D. Abraham, A. Blum, T. Sandholm, Clearing algorithms for barter exchange markets: Enabling nationwide kidney exchanges, in: Proc. of the ACM Conference on Electronic Commerce (EC), San Diego, California, 2007.

[2] J. Ahnand, J. Hornberger, Involving patients in the cadaveric kidney transplant allocation process: A decision-theoretic perspective, Management Science, 42 (1996) 629–641

[3] O. Alagoz,A. Schaefer, and M. Roberts, Optimization in Organ Allocation, 2009. In Handbook of Optimization in Medicine, P. Pardalos and E. Romeijn, Editors. Kluwer Academic Publishers.

[4] O. Alagoz, L. Maillart, A. Schaefer, and M. Roberts, Determining the acceptance of cadaveric livers using an implicit model of the waiting list, Operations Research 55 (2007) 24–36

[5] O. Alagoz, L. Maillart, A. Schaefer, and M. Roberts, The optimal timing of living- donor liver transplantation, Management Science, 50 (2004), 1420–1430

[6] I. Ashlagi, P Jaillet, V. H. Manshadi, Kidney Exchange in Dynamic Sparse Heterogenous Pools, Proceedings of the fourteenth ACM conference on Electronic commerce (2013) 25–26

[7] P. Awasthi, T. Sandholm, Online stochastic optimization in the large: Application to kidney exchange, In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)

[8] M. Beccuti, V. Fragnelli, G. Franceschinis, S. Villa, Dynamic simulations of kidney exchanges, in: Proceedings of International Conference on Operations research (OR2010), Springer, Munich, Germany, 2010.

[9] M. Beccuti, S. Haddad and G. Franceschinis, MDWNsolver: a framework to design and solve Markov Decision Petri Nets, International Journal of Performability Engineering, 7(5) (2011) 417428

[10] M. Beccuti, S. Haddad and G. Franceschinis, Markov decision Petri net and Markov decision well-formed net formalisms, Proceedings of the 28th international conference on Applications and theory of Petri nets and other models of concurrency (ATPN07),Siedlce, Poland, 2007, pp. 43–62.

[11] O. Boxma, I. David, D. Perry, and W. Stadje A new look at organ transplantation models and double matching queues, Probability in the Engineering and Informational Sciences 25 (2011), 135–155

[12] K. Cechlàrovà, T. Fleiner, D. Manlove, The kidney exchange game, in: SOR'05, L. Zadnik-Stirn and S. Drobne, eds., Slovenia, 2005, pp. 7–783.

[13] K. Cechlàrovà, V. Lacko, The kidney exchange problem: How hard is it to find a donor?, Annals of Operations Research, Article in press.

[14] Y. H. Chun, R. T. Sumichrast, A rank-based approach to the sequential selection and assignment problem, European Journal of Operational Research 174 (2006) 1338–1344.

[15] M. Constantino, X. Klimenova, A. Vianaa, and A. Raisd, New insights on integer-programming models for the kidney exchange problem, European Journal of Operational Research, 231 (2013), pp. 57–68.

[16] R. Davies, Study of a renal unit, Journal of the Operational Research Society 30 (1979) 873884.

[17] H. Davies, R. Davies, Simulation model for planning services for renal patients in europe, Journal of the Operational Research Society 38 (1987) 693–700.

[18] I. David, U. Yechiali, Time-dependent stopping probem with application to live organ transplants, Operations Research 33 (1985) 491–504.

[19] I. David and U. Yechiali, Sequential assignment match processes with arrivals of candidates and offers, Probability in the Engineering and Informational Sciences, 4 (1990) 413–430

[20] F. Delmonico, P. Morrissey, G. Lipkowitz, J. Stoff, J. Himmelfarb, W. Harmon, M. Pavlakis, H. Mah, J. Goguen, R. Luskin, E. Milford, G. Basadonna, M. Chobanian, B. Bouthot, M. Lorber, R. Rohrer, Donor kidney exchanges, Am J Transplant 4 (2004) 1628–1634.

[21] J. P. Dickerson, A. D. Procaccia, and T. Sandholm, Optimizing Kidney Exchange with Transplant Chains: Theory and Reality. Proc of the eleventh international conference on autonomous agents and multiagent systems, 2012.

[22] J. P. Dickerson, A. D. Procaccia, and T. Sandholm, Dynamic Matching via Weighted Myopia with Application to Kidney Exchange. Proc of the6th AAAI Conference on Artificial Intelligence, pp. 1340–1346, 2012.

[23] J. P. Dickerson, A. D. Procaccia, and T. Sandholm, Failure-aware kidney exchange. ACM Conference on Electronic Commerce, pp. 323–340, 2013.

[24] J. P. Dickerson, A. D. Procaccia, and T. Sandholm, Price of Fairness in Kidney Exchange, Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 2014.

[25] R. B. J. Freeman, Survival benefit: Quality versus quantity and trade-offs in developing new renal allocation systems, American Journal of Transplantation 7 (2007) 1043–1046.

[26] S. E. Gentry, D. L. Segev, R. A. Montgomery, A comparison of populations served by kidney paired donation and list paired donation. Am. J. Transplant , 5 (2005) pp.1914–1921.

[27] K. Glorie,B. Haase-Kromwijk, J. van de Klundert, A. Wagelmans, and W. Weimar Allocation and matching in kidney exchange programs, Transplant International, 27 (2014) pp. 333–343.

[28] J. Hornberger and J. Ahn, Deciding eligibility for transplantation when a donor kidney becomes available, Medical Decision Making, 17 (1997) 160–170

[29] D. Howard, Why do transplant surgeons turn down organs?: A model of the accept/reject decision, Journal of Health Economics, 21 (2002) 957–969

[30] L. Kallenberg, Markov Decision Processes, 2007, lecture notes.

[31] A. Kanematsu, N. Tanabe, K. ans Ishikawa, T. Tokumoto, S. Huchinoue, K. Takahashi, H. Toma, Impact of donor age on long-term graft survival in living donor kidney transplantation, Transplantation Proceedings 30 (7) (1998) 31183119.

[32] K. M. Keizer, M. de Klerk, B. Haase-Kromwijk, W. Weimar, The dutch algorithm for allocation in living donor kidney exchange, Transplantation Proceedings 37 (2005) 589–591.

[33] L. W. Kranenburg, T. Visak, W. Weimar, et al., Starting a crossover kidney transplantation program in the Netherlands: ethical and psychological considerations, Transplantation 78 (2) (2004) 194–197.

[34] M. Kurt, M. Roberts, A. Scheafer, and U.Ünver, Valuing prearranged paired kidney exchanges: A stochastic game approach. Boston College Department of Economics 785, (2011).

[35] A. W. Labib, An intelligent system for prioritisation of organ transplant patient waiting lists using fuzzy logic, Journal of the Operational Research Society 55 (2004) 103–115

[36] Y. J. Lee, J. H. Chang, H. N. Choi, et al., Donor-recipient age difference and graft survival in living donor kidney transplantation. Transplant Proceedings 44 (2012) 270-272.

[37] L. Lovász, M. D. Plummer, Matching theory, AMS Chelsea Publishing, Providence, RI, 2009, corrected reprint of the 1986 original

[38] R. M. Merion, V. B. Ashby, R. A. Wolfe, D. A. Distant, T. E. Hulbert-Shearon, R. A. Metzger, A. O. Ojo, F. K. Port, Deceased donor characteristics and the survival benefit of kidney transplantation, Journal of American Medical Association 294 (21) (2005) 2726-733.

[39] A. Nicolò, C. Rodriguez-Álvarez, Feasibility constraints and protective behavior in efficient kidney exchange, FEEM Working Paper 31.

[40] A. Nicolò, C. Rodriguez-Alvarez, Pairwise kidney exchange with age based preferences, in: ASSET2009-09-00275, Istanbul, Turkey, 2009.

[41] G. Opelz, Impact of hla compatibility on survival of kidney transplants from unrelated live donors, Transplantation 64 (10) (1997) 1473–1475.

[42] G. Opelz, Hla compatibility and kidney grafts from unrelated live donors, Transplantation Proceedings 30 (3) (1998) 704–705.

[43] F. Oppenheimer, P. Aljama, C. A. Peinado, The impact of donor age on the results of renal transplantation, Nephrology, Dialysis, Transplantation 19 (2004) iii11–iii15

[44] M. L. Puterman, Markov decision processes: discrete stochastic dynamic programming, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons Inc., New York, 1994, a Wiley-Interscience Publication.

[45] F. T. Rapaport, The case for a living emotionally related international kidney donor exchange registry, Transplantation Proc. 18 (1986) 59.

[46] R. Righter, Resource allocation problem in a random environment, Operations Research 37 (1989) 329–338.

[47] L. F. Ross, E. S. Woodle, Ethical issues in increasing living kidney donations by expanding kidney paired exchange programs, Transplantation LXIX (2002) 1539–1543.

[48] A. E. Roth, T. Sönmez, U. M. Ünver, Kidney exchange, The Quarterly Journal of Economics 119 (2) (2004) 457–488.

[49] A. E. Roth, T. Sönmez, U. M. Ünver, Pairwise kidney exchange, Journal of Economic Theory 125 (2) (2005) 151–188.

[50] A. E. Roth, T. Sönmez, U. M. Ünver, Efficient kidney exchange: Coincidence of wants in a structured market, American Econom. Rev. 97 (2007) 828–851.

[51] D. L. Segev, S. E. Gentry, D. S. Warren, B. Reeb, R. A. Montgomery, Kidney paired donation and optimizing the use of live donor organs, JAMA 293 (2005) 1883–1890.

[52] D. L. Segev, S. E. Gentry, J. K. Melancon, R. A. Montgomery, Characterization of waiting times in a simulation of kidney paired donation, American Journal of Transplantation 5 (2005) 2448–2455.

[53] X. Su, S. A. Zenios, Patient choice in kidney allocation: A sequential stochastic assignment model, Operations Research 53 (2005) 443–455.

[54] U. Ünver, Dynamic kidney exchange, Review of Economic studies 77 (1) (2009) 372–414.

[55] S. Villa, F. Patrone, Incentive compatibility in kidney exchange problems, Health Care Manage. Sci. 12(4) (2009) 351–62.

[56] Y. Yuan, S. Feldhamer, F. Fyfe, D. Ludwin, The development and evaluation of a fuzzy logic expert system for renal transplantation assignment: Is this a useful tool?, European Journal of Operational Research 142 (2002) 152–173.

[57] S. A. Zenios, E. S. Woodle, L. F. Ross, Primum non nocere: avoiding harm to vulnerable wait list candidates in an indirect kidney exchange, Transplantation 72 (4) (2001) 648–654.

[58] S. A. Zenios, Optimal control of a paired-kidney exchange program, Manage. Sci. 48 (3) (2002) 328–342.

[59] S. A. Zenios, L. M. Chertow, G. M.and Wein, Dynamic allocation of kidneys to candidates on the transplant waiting list, Operations Research 48 (2000) 549569.