OPTIMAL
RED BLOOD CELL
MATCHING

Joost van Sambeeck

OPTIMAL RED BLOOD CELL MATCHING

Joost van Sambeeck

Promotie commissie

Voorzitter & secretaris:	Prof. dr. J.N. Kok
	University of Twente, Enschede, the Netherlands
Promotors:	Prof. dr. N.M. van Dijk
	University of Twente, Enschede, the Netherlands
	Prof. dr. W.L.A.M. de Kort
	University of Amsterdam, Amsterdam, the Netherlands
	Dr. ir. M.P. Janssen
	Sanquin Research, Amsterdam, the Netherlands
Leden:	Prof. dr. J.T. Blake
	Dalhousie University, Halifax, Canada
	Prof. dr. R.J. Boucherie
	University of Twente, Enschede, the Netherlands
	Prof. dr. M. de Haas
	Leiden University, Leiden, the Netherlands
	Prof. dr. ir. E.W. Hans
	University of Twente, Enschede, the Netherlands
	Prof. dr. ir. D. den Hertog
	University of Amsterdam, Amsterdam, the Netherlands
	Prof. dr. C.E. van der Schoot
	University of Amsterdam, Amsterdam, the Netherlands

OPTIMAL RED BLOOD CELL MATCHING

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
Prof. dr. ir. A. Veldkamp,
volgens besluit van het College voor Promoties,
in het openbaar te verdedigen
op woensdag 9 december 2020 om 16:45 uur

door

Josephus Henricus Jacobus van Sambeeck

geboren op 28 september 1991
te Reusel, Nederland

Dit proefschrift is goedgekeurd door:

Prof. dr. N.M. van Dijk
Prof. dr. W.L.A.M. de Kort
Dr. ir. M.P. Janssen

Contents

Part I

Introduction

Introduction

A blood transfusion is a safe, common, and potentially life-saving medical procedure in which one or multiple blood components, originating from a donor, are inserted into the bloodstream of a transfusion recipient. One of these blood components is the red blood cell, by 400,000 units per year in the Netherlands [99] and 85 million units per year worldwide [129] the most frequently transfused blood component. Red blood cell transfusions are typically used to improve the oxygen-carrying capacity of the blood. The recipient might, for example, suffer from a genetic disorder that affects the functioning of the red blood cells (sickle cell disease, thalassemia), cancer or a cancer treatment that affects the red blood cell production (leukaemia, chemotherapy, stem cell transplant), or severe bleeding (surgery, childbirth, trauma).

With respect to transfusion of red blood cells, it is important that the blood groups of the donor and transfusion recipient match. The blood group of an individual is determined by the presence or absence of antigens on the surface of the red blood cells. If a particular antigen is present on red blood cells of a donor but absent on the red blood cells of the transfusion recipient, the immune system of the transfusion recipient may produce antibodies against this foreign antigen. This is called alloimmunization. These antibodies will cause problems during a subsequent transfusion or, in the case of a female recipient, a future pregnancy. Such a response from the immune system can be prevented by selecting red blood cells lacking the relevant antigen.

1.1 Red blood cell matching

1.1.1 History

The first successful human-to-human blood transfusion has been reported in 1825 by James Blundell, an obstetrician at Guy's and St. Thomas' Hospitals in London. He treated a woman with severe postpartum hemorrhage (excessive blood loss after

childbirth). The woman received four ounces of blood from her husband, rallied, and survived. Between 1818 and 1830 Blundell performed ten blood transfusions with varying degrees of success. He acknowledged that there were serious risks associated with his procedure and therefore only applied it in exceptional cases [5, 11]. At the end of 19th century, blood transfusion was regarded as a dangerous procedure due to its varying degrees of success. It was, therefore, avoided.

Below a brief historic description of the originating blood group systems is provided, as adopted from [48]. In 1901, Karl Landsteiner discovered the ABO blood groups. He described the reactions between the red blood cells and sera of 22 individuals [68] and observed that the addition of serum from some individuals reacted with the red blood cells of other individuals. He recognized a pattern and showed that the individuals could be divided into three blood groups: A, B, and O. For example, the serum of individuals belonging to blood group O reacted with the red blood cells from groups A and B. In the next year, two of his students confirmed his findings in a larger study among 155 individuals. They also found four subjects (2.5%) that did not belong to one of these three groups. Later this became the fourth blood group, blood group AB.

A quarter of a century later, another student of Landsteiner encountered a problem with a transfusion between a man and woman, who had both blood group O. [71]. Combining the woman's serum with cells from her husband resulted in a reaction. Her serum was also combined with the red blood cells of 104 other individuals with blood group O and a reaction was seen in 80 cases. This was the first discovery of the Rhesus-D antigen. The name for this blood group system came from parallel experimental work carried out by Landsteiner and Wiener involving research on rabbits and guinea pigs with blood from Rhesus monkeys. The serum from these animals was also found to react with the red blood cells of 85% of the individuals tested, who were classified as Rhesus positive [48].

This work stimulated similar research and many other antigens were recognized in subsequent years. The identification of new blood group systems was facilitated by the development of the anti-globulin test as well as the recognition that incubation of red blood cells with enzymes enhanced the expression of some antigens [24, 76]. The Kell system was the first blood group system that was identified through the application of this test [25]. The discovery of the Duffy (Fy) and Kidd (Jk) systems followed quickly.

1.1.2   Present

Currently there are more than 300 antigens known, of which only 25 are clinically relevant [108]. In practice, however, extensively typed blood products are only applied for specific groups of transfusion recipients. These groups consist of recipients who are expected to receive multiple transfusions or recipient who have an increased risk of developing antibodies (e.g., patients who have developed an-

tibodies in the past). For example, recipients with sickle cell disease, thalassemia, auto-immune hemolytic anemia (AIHA), or myelodysplastic syndrome (MDS), recipients who already have developed antibodies, women of reproductive age (< 45 years) [18]. All individuals who have developed antibodies (as a result of a previous blood transfusion) are logged in national register (TRIX). Previous to a new transfusion this register is checked for presence of antibodies, which might no longer be detectable in the patients blood.

To be able to comply with the Dutch transfusion policy, Sanquin has an extensively typed donor base [125]. This implies that the majority of the donors are typed for the ABO, Rhesus and Kell blood group systems. Dependent on the blood group profile of the donor on this limited number of antigens, the donor is further tested for other blood group systems. The blood group of the donor is determined by serology. This implies that for each antigen a separate test has to be performed, which is both costly and time-consuming. Although a large part of the donor population has been extensively typed, hospitals still face the problem of ensuring that patients who have developed antibodies will obtain matched blood products.

1.1.3   Future

Recent technological developments in diagnostics enable blood group identification by genotyping instead of serology. This implies that with one single measurement the presence or absence of all antigens (or at least more than a hundred) can be determined. This creates a situation in which it becomes practically feasible to determine the extended blood group of donors and recipients. In addition, the costs of genotyping itself are expected to become comparable to the costs of a single serological test in the forseeable future. These developments will thus steer towards the application of extended blood group matching for all transfusion recipients. This implies a paradigm shift from preventing the consequences of alloimmunization to the prevention of alloimmunization itself by extended matching. One of the additional benefits from extended matching is that the in-hospital testing for antibodies can be abolished. The magnitude of the beneficial impact of red cell genotyping on patient outcomes compared to traditional serology-based laboratory methods, remains unclear and warrants further study [15].

1.2   Managing the blood supply

The management of the inventory of blood products has a number of specific features. These are 1) compatibility of blood components and 2) perishability of blood components. In an ideal situation, one would transfuse only identical blood products. However, due to the compatibility of blood groups, some blood products can be used to satisfy multiple requests, as long as these are compatible. So, there

are multiple differing inventories (one per blood group) which can satisfy different requests. Ideally, one would have only one inventory with blood groups that are negative for all antigens and therefore could satisfy any specific blood group request. This unbalanced/unequal usability of the blood groups in the inventory creates a very specific inventory management approach. In addition to the unbalanced usage of blood groups, blood inventory management requires taking into account the fact that red blood cells outdate after 35 days.

When issuing a blood product from inventory, one has to consider both of the features mentioned above (compatibility between the supplied and requested blood groups and their perishability) in such a way that not only the current request can be satisfied, but that the remaining inventory will allow doing so for any future request as well. For the basic ABO-D matching this balance is overseeable as there will always be an identical red blood cells available (due to the limited number of blood groups). However, when the number of antigens considered increases, the complexity of the inventory increases as well, which will force an increase in the number of incompletely matched blood groups.

1.2.1 Managerial decisions

The development of inventory models for perishable products dates back to the 1960s. Since that time several review papers have been written with applications in the fields of food, pharmaceuticals, photographic films, drugs, and blood. Although there are no review papers that specifically focus on red blood cell matching, it is often considered a sub-part of the inventory management of blood products, which is an area in OR that has extensively be studied. For an overview for review on either perishable products or blood products we refer the reader to the following review articles:

- perishable inventory management [4, 50, 62, 78, 90]

- blood inventory management [7, 83, 85, 86, 88, 107]

When we have to decide which red blood cell units are issued from inventory there are two main decision that have to be made:

- What is the blood group of the issued red blood cell unit?

- What is the age of the issued red blood cell unit?

While most of the papers on perishable inventory have assumed a fixed lifetime of the products, few papers have also considered the life to be a random variable, which is often exponentially distributed. We assume that red blood cell units have a fixed lifetime of 35 days and that incoming units have age zero when they enter the inventory.

1.3 Contributions and thesis outline

In contrast to existing literature on inventory management of perishable products, or more specifically, blood products, the main focus in this thesis is on inventory issuing policies when extended blood group matching would be applied. This implies that the models presented can be applied for both the current setting, where the blood groups of individuals are determined by serology, as well as for a future setting where mass-scale genotyping is implemented.

Part I: Introduction

The introduction consists of this chapter and Chapter 2. In **Chapter 2** we address the concept of alloimmunization, the most frequent adverse event of blood transfusions. Whilst completely matched donor blood would nullify the alloimmunization risk, this is practically infeasible. Current matching strategies, therefore, aim at matching a limited number of blood group antigens only, and have evolved over time by systematically including matching strategies for those antigens for which (serious) alloimmunization complications most frequently occurred. An optimal matching strategy for controlling the risk of alloimmunization, however, would balance alloimmunization complications and costs within the entire blood supply chain, whilst fulfilling all practical requirements and limitations.

Part II: Donor selection

In **Chapter 3**, the objective is to compute the effectiveness for recruiting next of kin's for donorship. For rare blood groups the recruitment of donor relatives, for example siblings, is expected to be effective, since the probability of a similar rare blood group is likely. However, the likelihoods strongly differ between blood groups and are not commonly available. This chapter provides a unified mathematical formulation to calculate such likelihoods. From a mathematical and probabilistic point of view, it is shown that these likelihoods can be obtained from the computation of a stationary genotype distribution. This, in turn, can be brought down to a system of quadratic stochastic operators. A generic mathematical approach is presented, which directly leads to a stationary genotype distribution for arbitrary blood groups. The approach enables an exact computation for the effectiveness of recruiting next of kin for blood donorship. Next to an illustration of computations for ABO and Rhesus-D blood groups, it is particularly illustrated for the extended Rhesus blood group system. Other applications requiring next of kin blood group associations can also be solved directly by using the unified mathematical formulation.

Part III: Inventory allocation

Chapter 4 is a preliminary chapter for **Chapters 5**, **6** and **7** In this chapter we introduce a binary representation for general blood groups (i.e. beyond ABO,

Rhesus-D blood groups). This binary representation provides a clear and unambiguous way to represent blood groups mathematically, irrespective of the number of antigens considered. Another advantage of this binary representation is that a compatibility matrix, which grows exponentially with the number of antigens considered, is no longer required, since the compatibility between blood groups can be easily determined by an element-wise comparison of binary vectors. In addition, it presents how the blood group distribution can be computed from phenotype frequency tables for an arbitrary set of antigens.

In **Chapter 5**, we model the inventory allocation problem as a Markov Decision Process. First we show that the deterministic ageing process of an red blood cell unit can be approximated by a phase-type distribution, or more specifically, an Erlang distribution. We then provide a generic description and formulation for issuing red blood cell units upon requests. Its detailed description can be used to further develop approximative techniques for solving.

The Markov Decision Process formulation in **Chapter 5** is computationally very demanding and can only be solved for small problem instances. In **Chapter 6** we therefore apply a decision rule, which is based on both the age and rareness of the red blood cell units in inventory. More specifically, we compute the relative opportunity loss between the blood group of the RBC units in inventory and the RBC units requested. Using this predetermined decision rule, the inventory allocation problem can be modelled as a minimal cost flow problem, which can easily be solved, also for realistic problem instances consisting of up to 2^{14} different blood groups.

In **Chapters 5** and **6** we incurred a shortage, when there was an insufficient number of matching red blood cell units available from inventory. In practice, however, not being able to satisfy a request is unacceptable and a blood product must be issued. The solution is to issue a red blood cell unit with a smaller number of matching antigens. In **Chapter 7** we deal with this issue and determine an optimal order for antigen exclusion when an insufficient number of matched RBC units is available from inventory, such that the probability that antibodies will be developed is minimized.

Finally, in **Chapter 8**, we give some general conclusions and provide directions for future research based on the findings in this thesis.

A conceptual framework for optimizing blood matching strategies

2.1 Introduction

In a utopian world every blood transfusion would be handled like an organ transplant, which means that one would try to find a perfect match between donor and recipient. The reality however is that completely matched donor blood is impossible in practice due to the abundance of blood group antigens, costs associated with blood typing, and complications the logistics for such a scheme would impose. As a consequence only a handful of blood group antigens are matched, placing transfusion recipients at risk for alloimmunization and associated transfusion complications. An ideal matching strategy would be one that minimizes the risk of alloimmunization, is cost-effective, and fits within the practical limitations of the blood supply chain. In the past, matching strategies have been guided by the frequency of alloimmunization incidents, without systematically considering all consequences such strategies impose on the blood supply. Since a selected matching strategy will either directly or indirectly affect the entire blood supply chain (Figure 2.1), an integrated approach is required. Such an approach would, for any particular blood matching strategy, allow balancing the costs of donor recruitment, donor typing, inventory management, blood product logistics, patient blood typing, and alloimmunization complications in transfusion recipients. Besides costs also the effects of transfusion complications on patients' health should be taken into account. This article describes the outline of a generic integrated blood management model, its components, their interaction and potential complicating factors and limitations currently foreseen for such a model.

We will first provide a description of all elements within the blood transfusion chain that are relevant to such a blood management model. Next we will describe how various elements are combined into an integrated model. Finally, we will discuss which challenges are foreseen with the implementation of the model and

Matching

Antigens in patient

Donor

Stock

Patient

1

2

1. Depending on the demand of blood types and the distribution of blood products currently available in stock, donors with particular blood types are needed.

2. The requirement of particular blood types directly affects the availability of blood products in stock. This is therefore directly affected by the patient's characteristics (blood type and blood use) in combination with the matching strategy applied.

Figure 2.1: Schematic overview of blood type matching and its impact on the blood supply chain.

potential prospects. Challenges will concern knowledge required for shaping the modeling structure and the availability of data for various model parameters. Not only will the model guide the search for a rational choice of an optimal matching strategy, it will create transparency for the decision arena: the balance between costs and patient outcomes will become explicit for whatever optimal decision is selected. Secondly, by developing an integrated model, any blind spots in knowledge regarding any of the elements of the decision model will become visible.

The elements identified for the integrated blood management model are: the patient population, transfusion practice, pre-disposition of transfusion complications, typing and matching strategies, and the donor population. Note that as the patient is the primary concern, it is the patient that should be the starting point of the analysis. From there we will work our way back through the blood transfusion chain towards the donor population.

2.2 Transfused patients, exposure and transfusion complications

Blood transfusion is a common medical procedures performed in hospitals. Despite its benefits, patients exposed to red blood cell (RBC) antigens may produce antibodies, which can cause acute or delayed hemolytic transfusion reactions (HTR). In addition, upon pregnancy in alloimmunized women, hemolytic disease of the fetus and new-born (HDFN) may occur. Not all patients form antibodies after RBC transfusion. According to current views, most are so-called 'non-responders' and will never form antibodies despite numerous transfusions. Others seem to have an increased immunization risk and develop multiple antibodies after a few antigenic exposures, these are referred to as the '(hyper)responders' [46]. It is currently not possible to prospectively identify patients that will form antibodies. In the absence of phenotypic matching, RBC alloimmunization risks vary between patient groups; it occurs in less than 5% of all transfusion recipients, increases to about 10 - 30% in patients with thalassemia, auto-immune hemolytic anemia or myelodysplastic syndromes, and can be more than 50% in sickle cell anemia patients [123, 124]. In addition, patients with antibodies are at increased risk for additional antibody development upon subsequent transfusions [59, 103]. During pregnancy, maternal RBC antibodies against paternal inherited antigens can pose the child at risk for HDFN. Besides anti-D, anti-E, anti-K, and anti-c are the most frequently encountered antibodies with the potential to seriously complicate pregnancy if the fetus carries the cognate antigen. The risk for severe HDFN in these fetuses, requiring intra-uterine or postnatal (exchange) transfusion, is estimated to be 12% for anti-K, 8.5% for anti-c and about 1% for anti-E. While for anti-D, administration of anti-D immunoglobulin (besides preventive D-matching) has reduced the risk of D immunization from 15 to 0.3 percent, such measures are not available or not always applied for other antigens, which are in the majority of cases elicited by previous transfusions [66].

The impact of transfusion reactions may vary widely, ranging from serologic observations or mild symptomatic anemia only, to life-threatening complications and death. It is obvious that with increasing severity, costs of treatment will also increase, although studies reporting on such associations and associated costs are currently limited or completely lacking [63]. Maximum benefits of alloimmunization prevention can be obtained by administering extended antigen matched blood to patients who have an a priori high risk for alloimmunization. Therefore, unravelling genetic and environmental conditions enhancing RBC immunization would support preventive strategies. Although most studies on this subject have been performed in sickle cell disease (SCD) patients, factors such as age, sex, inflammatory status, MHC class-II genotype, polymorphisms associated with immune modulation and altered immune (regulatory) cells and disease or therapy associated immuno-

suppression seem to influence the immune response towards transfusion exposed alloantigens [6, 46, 67, 73, 109, 130, 134]. Due to logistic constraints, elaborate preventive matching based on a responder-profile is expected to be only feasible for a small proportion of patients. Targeting patients with (chronic) elective transfusions is likely to be feasible. Also, two recent prospective studies showed that less than 50% of surgery patients, who according to the local hospital pre-operative blood-ordering schedule had a high transfusion risk, were actually transfused. Extensive preventive matching as a routine policy is therefore expected to require a substantial amount of additional work and costs. Moreover, about 25% of patients required more than the anticipated number of RBC units during surgery and extended matched units were not readily available [34, 104].

As the blood management model is aiming to optimize strategies for preventing HTRs, the risk of alloimmunization in patients, its associated cost and health impact needs to be defined. The ongoing Dutch R-fact study –in which the predisposition for formation of antibodies is studied– will allow modelling the likelihood of antibody formation. This information, combined with data on blood use for various patient groups, which will be obtained from the Dutch PROTON study (in which detailed transfusion data from a large number of hospitals are combined in a Dutch Transfusion Datawarehouse), will provide the information required to model the likelihood of HTRs in various patient groups. Research on the cost and health impact associated with HTRs will also be required to complete the model for patient and health outcome of transfusion complications.

2.3 Current matching strategies in the Netherlands

In the Netherlands all RBC transfusions are compatible for ABO and D antigens. Since 2011 the guideline for selection of RBC units prescribes preventive matching for specific blood group antigens for different patient subgroups. Since 2004 it has been policy to select K-negative RBCs for women aged under 45, which in 2011 was extended with matching for c and E. These measures aim to prevent HDFN. In the updated guideline four patients groups with a putative increased risk of alloimmunization were defined, on grounds of either underlying disease, transfusion frequency, or potential (hyper-)respondership. The four patient groups concern 1) patients with autoimmune hemolytic disease; 2) patients with myelodysplastic syndrome and 3) patients with an immediate early antibody (IEA) against a clinically relevant RBC antigen. For these three patient subgroups Rh phenotype (CcDEe) and K compatible RBCs are selected. Finally, the fourth group consists of patients with hemoglobinopathies (SCD or thalassemia) for whom Rh phenotype, K and Fy^a compatible RBCs are selected, and whenever available, Jk^b, S or s compatible RBCs. The recommended matching strategies formulated in Dutch transfusion guidelines are summarized in Table 2.1 [18].

Table 2.1: Matching strategies for various patient groups as recommended in the 2011 Dutch Transfusion guideline.

Patient group	Matching strategy
Sickle cell anemia and thalassemia	Rh phenotype, K and Fy[a] (and if available, Jk[b], S and s)
Autoimmune hemolytic anemia	Rh phenotype and K
Myelodysplastic syndrome	Rh phenotype and K
Alloimmunized with clinical important antibodies	Rh phenotype and K
Woman of childbearing age	c, E and K

Apart from these specific patient groups, patients in the Netherlands are routinely tested for the presence of IEAs prior to RBC transfusions. When IEAs are detected, both their specificity and clinical importance are investigated. In case of a clinical important IEAs it is essential to select donor erythrocytes that are negative for corresponding antigens to prevent HTRs. Furthermore, dependent on the matching strategy, it may be required that donor erythrocytes are compatible with other antigens of the patient (extended matched), to prevent the formation of additional IEAs. Because antibodies may lose detectability over time, accurate recording and accessibility of patient antibody formation is of the utmost importance [93, 95, 102]. Besides in-hospital records, a national database is available in the Netherlands (TRIX, Transfusion Register Irregular antibodies and X(cross)-matching), in which hospitals register patients with RBC antibodies and cross-match problems [120]. This system is accessed for the evanesced antibodies in all patients with a transfusion request to prevent re-exposure to the cognate antigen. However, these registrations will not prevent re-exposure due to an inadequate antibody follow-up after transfusion.

The blood management model will have to accommodate matching strategies currently implemented as well as various extended matching strategies. The model should incorporate all costs involved for various matching strategies considered (e.g. costs of personnel and materials used).

2.4 Typing the donor population

Different matching strategies will pose different requirements on the availability of typed blood products. The required number of typed blood products, the variation in its demand, and the required service level (the probability of not being able to deliver a requested typed blood product) will determine the number of typed blood products that will have to be available in stock at any time, and hence the level of typed donors. A large typed donor population has the advantage that in

most cases donor erythrocytes can be selected directly from inventory, even when blood products need to be typed negative for combinations of antigens. However, there will always be a balance between the additional efforts required to fulfill requirements for typed blood products and extending the pool of elaborately typed donors.

2.5 Donor recruitment

Transfusing matched blood is only feasible if there are enough donors that are typed negative for specific (combinations of) blood group antigens. For instance, many Blood Services in Western countries have a structural shortage of Fy^a-neg, Fy^b-neg, e-neg donors. This blood type is most common in populations from Sub-Saharan Africa, of which relatively few individuals are enrolled as blood donors [119]. In addition, in many countries a broad variety of ethnic minority populations exist. Shifting immigration patterns and mixing of these populations will increase the demand for rare blood type combinations. A valuable 'side effect' of recruiting among minority groups is a potentially increase of donors for HLA-matched substances of human origin, such as stem cells. Blood Services therefore need to identify which specific ethnic minority populations to focus on in terms of rare blood type prevalence.

2.6 Integration

In the previous sections various elements of the blood transfusion chain and their interdependencies were discussed (see Figure 2.1). Each of these elements and their interactions need to be modelled to allow evaluation of the impact of a particular matching strategy on the transfusion risk of patients (i.e., acute and delayed HTRs) and on other parts of the blood supply chain (e.g., the availability of matched blood products, costs of type and screen, storage, outdating, and targeted donor recruitment). The main elements of the blood supply chain and the associated sub-models describing various interactions required for an integrated blood management model is depicted in Figure 2.2.

The starting point for any evaluation is the blood matching strategy, as this, in combination with the patient mix, will determine the demand for particular blood products. Depending on the matching strategy and patient mix (patient subgroups) there will be a risk of antibody formation and subsequent risk for adverse transfusion complications. Moreover, the combination of patient mix and associated matching strategy will determine the demand for typed blood products in the inventory. The availability of typed blood products in the inventory is dependent on the availability of typed blood donors, which again is dependent on the efforts and requirements of targeted donor recruitment.
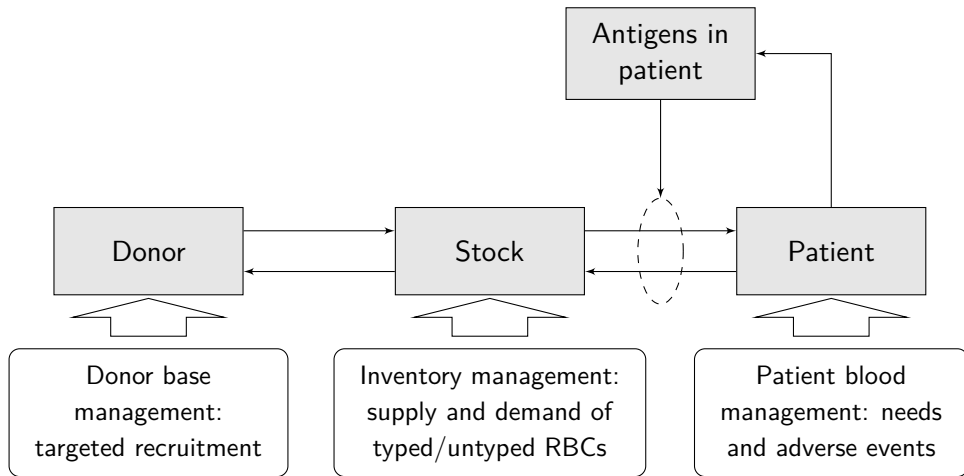
Figure 2.2: Main elements of the blood supply chain and associated sub-models of the blood management system.

The assessment of the transfusion complication risk requires estimates of the likelihood of antibody formation and subsequent transfusion reactions in patients given a particular matching strategy. Such estimates should incorporate the transfusion pattern and the ethnic (blood type) composition of various patient subgroups. Also, antigen specific estimates for the likelihood of developing antibodies as well as for transfusion complications are required. The likelihood of transfusion complications in combination with cost and the health impact will allow estimation and subsequent balancing of the costs and benefits from the matching strategy applied.

To enable matching blood for transfusion recipients antigen and antibody profiles of patient subgroups have to be determined. Next, compatible RBC units have to be selected from inventory. Detailed information on blood use and the antigen profiles per patient group allows assessment of the blood inventory required to meet patient needs. This will be a description of the required inventory both in terms of amount and composition of RBCs in various stocks along the blood transfusion chain. Blood product demand will show a stochastic behavior and a realistic blood management model will therefore have to be able to accommodate such random variations. Given the patient mix, matching strategy and associated transfusion characteristics, for any pre-specified acceptability rate for the unavailability of (matched) blood products and inventory management strategy, the required blood inventory size and composition can be determined. The resulting costs and effects for the complete blood transfusion chain (outdating, size of the inventory, logistics, and material handling costs) can now be estimated. Note that the unavailability of matched blood products will impact the likelihood of transfusion complications in patients. Therefore, optimization of the overall

blood transfusion chain will require a separate sub-optimization for the inventory management strategy.

The availability of compatible RBC units required in the inventory is directly linked to the availability of typed donors and hence guides the typing strategy and targeted donor recruitment efforts. The typing strategy will be aiming at fulfilling the requirements for maintaining sufficient inventory levels, but this will be dependent on the availability of specific antigen profiles in the (typed) donor population. Whenever these are insufficient, targeted donor recruitment efforts will have to ensure adequacy of the desired antigen profiles in the un-typed donor population, and ultimately those in the typed donor population. Estimates for the costs of recruiting specific donor subgroups in order to ensure a sufficient level of typed blood groups in the donor population are required to estimate the costs for maintaining the required inventory levels. Other than in the inventory management, which is an in-line process, it is presumed that the required levels of typed donors will be met by increasing donor recruitment efforts.

2.7 Discussion

In this article we discussed a conceptual framework for a blood management model which allows optimization of blood matching strategies. The model links various elements from the blood transfusion chain to allow an assessment of the full impact of any particular matching strategy. The approach is unique in the sense that in the past matching strategies were guided by the prevention of transfusions complications observed with the administration of blood products, without consideration its impact on the underlying blood supply process. In theory this new approach seems sensible, however, in practice there will be a number of complicating factors.

First of all, except for some specific patient subgroups there is only limited evidence available on the effectiveness of matching strategies for the prevention of transfusion complications. Despite the fact that transfusion complications are accurately analyzed, patient exposure is far more difficult to ascertain. More evidence however has been gained for the risks of alloimmunization in various patient cohorts in the Netherlands in the ongoing Risk-Factors for alloimmunization after red blood Cell Transfusion (R-FACT) study [133]. This concerted collaboration of several large hospitals will provide the information required to model risk factors for some patient subgroups. Also, looking back at the reduction of transfusion complications after implementation of altered matching strategies may support inference on its effectiveness. However, this effect may also be confounded by transfusion practice.

Another complicating factor is the impact of transfusion complications on patients, as this may vary from serologic observations or mild symptomatic anemia to life-threatening complications and death. Not only are predictors for predisposing

factors lacking, but the impact of various levels of transfusion complications on patient health (apart from death) are not readily available, and neither are the associated costs. Assessing costs of complications is complex as it requires separation of the costs of patient treatment from costs of complications which are confounded by definition. Similar complications occur when estimating the impact on patient health. Nonetheless, an increasing number of publications on the impact of transfusion complications are becoming available [47, 81, 100].

In most settings detailed information on transfusion practice (number of transfused blood products for specific patient subgroups and the variation herein) is lacking. In the PROTON II study for a large number of Dutch hospitals detailed information on blood transfusions administered to patients is collected in one central data warehouse [121]. These data consist not only of transfused products, but also patient diagnosis and lab results. These data are indispensable when modelling the logistics of the blood supply in general, and for specific patient groups. Optimized inventory and dispatching strategies can be developed for both hospital and regional distribution centers and may be tailored to specified matching strategies. Note that with data on blood use the requirements and constraints for such models are available.

For the assessment of the risk of transfusion reactions (depending on the matching strategy) information on historical exposure of patients to blood products is required in order to assess the likelihood of antibody development. Such data is at present only available at a large scale for Denmark and Sweden where long term follow-up data on transfused patients is recorded in the SCANDAT database [32, 33]. Such information may be used to estimate an approximate risk of exposure to red blood cells in other settings.

The development of an integrated blood management model will increase transparency in costs and effects of selected matching strategies and is therefore -if applied- expected to contribute to an improved efficiency in blood transfusion practice.

Part II

Donor selection

CHAPTER 3

Blood group probabilities by next of kin

3.1   Introduction

3.1.1   Motivation

The challenges faced by blood transfusion services are becoming more complex and are changing continuously due to growing economic pressure, new technologies, and increasing customer expectations [26, 127]. One of these expectations is the ability to select extensively (blood group) matched red blood cells (RBC's) for transfusion recipients, to decrease the number and severity of transfusion reactions. However, current blood donor recruitment strategies are based on historical matching strategies and cannot meet the demand for extensively matched blood products. Furthermore, due to increasing immigration rates and differences in blood group distributions between ethnic populations the diversity among blood groups within the transfusion population increases. For instance, the blood group profiles of Caucasian individuals (i.e., individuals with European ancestors) and individuals from African descent differ significantly. In contrast, in the donor base ethnic minorities are underrepresented, complicating extended blood group matching of donors and transfusion recipients. Hence, one of the major challenges for current blood donor recruitment practice is to maintain an adequate donor base with a sufficiently diverse blood group composition [16]. In actual fact, an overrepresentation of donors from African descent would be preferable, as individuals from African descent have a higher probability of requiring repeated blood transfusions as a result of sickle cell decease, which is uncommon in other populations [14].

In practice, it has been shown to be effective to increase the number of donors with O, Rhesus-D (RhD) negative blood groups by recruiting among their relatives, since these are more likely to be O, RhD-negative than individuals in the general population. Although intuitively this seems to be an effective strategy, it is not evident to what extent such strategies are more effective than random donor selection. Moreover, it gives rise to the question whether this also holds for other blood

group combinations. If so, it may inform more effective recruitment strategies.

3.1.2   Approach

To model blood group antigen inheritance quadratic stochastic operators (QSO's) are used, as introduced by Bernstein in 1924 [8]. Recently, Ganikhodjaev et al. [42, 43, 44] applied QSO's to model the heredity of ABO and RhD blood groups. However, a general formulation that goes beyond the standard ABO, RhD blood groups was not given. In addition, an exact computation of the effectiveness of recruiting relatives of donors with rare blood groups has not been included. Of course, the idea that relatives have similar blood groups is intuitively correct, but quantification is insightful and allows balancing recruitment efforts against the benefits from blood group matching.

This chapter presents a unified mathematical formulation to determine the probability that two relatives (next of kins) share the same blood group. In short, the steps and formulation that will be provided, transform phenotype distributions into genotype distributions, and back. By this generic mathematical approach we can directly analyse the effectiveness of specific next of kin recruitment strategies, for any blood group, ethnicity, and population (as numbers may differ worldwide).

The mathematical approach only requires a phenotype distribution as an input, whereas the population genotype distribution is required for calculating the blood group distribution probability for the next of kin. Phenotype distributions can be easily determined by simple blood tests, genotype distributions are more difficult to obtain. However, these genotype distribution can be derived from the phenotype distributions using our generic mathematical approach.

This chapter is organized as follows. Section 3.2 starts with a known, but motivational example for the ABO, RhD blood groups. Next, in Section 3.3, a unified mathematical formulation of the approach is covered. In section 3.4 this unified mathematical formulation is used the compute the effectiveness of recruiting next of kin for blood donorship. Finally, we explore some specific applications of the approach in Section 3.5. Section 3.7 (Appendix I) provides a clear overview of the notation used.

3.2   Motivational and illustrative example

In this section, let us first provide the genetic backgrounds and illustrate our steps and formulation for the 'standard' ABO, RhD blood groups. That is, we show

- how the approach for determining the distribution of genotypes in a population essentially comes down to a system of quadratic equations,

- how the distribution of genotypes can be used to evaluate the effectiveness of targeted recruitment strategies for the ABO and RhD blood groups

separately,

- how the results for both blood groups can be combined.

Later, in Section 3.3 and 3.4, the same steps and approach are provided in a unified mathematical formulation, such that this formulation can be applied to any blood group system.

3.2.1 ABO, RhD blood groups

According to the International Society of Blood Transfusion (ISBT) there are more than 300 different blood group antigens belonging to 36 blood group systems [108]. Each antigen can be either present or absent on the surface of an RBC, leading to an extremely large number of different blood group profiles. In practice, however, not all antigens are equally important with regard to transfusion related problems. The most important antigens are A and B (both belonging to the ABO blood group system), followed by RhD, which belongs to the Rhesus (Rh) blood group system. Taking only these three antigens into consideration the total number of blood group profiles can be compressed into eight major groups, the so-called ABO, RhD blood groups. These ABO, RhD blood groups consist of a combination of a blood group belonging to the ABO blood group system (O, A, B, AB) and a RhD blood group (RhD-neg (d), RhD-pos (D)).

For just the RhD blood groups three different genotypes ($\mathcal{G}_D = \{dd, Dd, DD\}$) and two different phenotypes ($\mathcal{F}_D = \{d, D\}$) exist, where the genotype is a genetic code that determines which antigen might be expressed on the surface of the RBC's. The expression of particular antigen is called the phenotype. Moreover, multiple genotypes may lead to the same phenotype. The relation between the different RhD genotypes and phenotypes is shown in the following matrix:

$$S = \begin{array}{c} \\ dd \\ Dd \\ DD \end{array} \begin{array}{cc} d & D \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \end{array}, \tag{3.1}$$

where a one indicates which genotypes results in a particular phenotype. Note that genotypes (and genes) are presented in italics and phenotypes (and antigens) are presented in a regular typeface.

Similarly, the ABO blood group system consists of six different genotypes ($\mathcal{G}_{ABO} = \{OO, OA, OB, AA, AB, BB\}$) and four different phenotypes ($\mathcal{F}_{ABO} = \{O, A, B, AB\}$). The relative frequencies for the ABO, RhD blood groups in the general Caucasian population are given in Table 3.1. The RhD and ABO blood groups belong to two blood group systems and are inherited independently. Therefore, in the next sections we will explore which steps are required to investigate

Table 3.1: Relative frequencies for the ABO, RhD blood groups [94].

		O	A	B	AB
		0.44	0.43	0.09	0.04
d	0.168	0.074	0.072	0.015	0.007
D	0.832	0.366	0.358	0.075	0.033

the effectiveness of recruiting next of kin with respect to the RhD and ABO blood groups separately. At the end of this section the results for both blood groups are combined.

Note that most of the computations performed in this section are similar to what can be found in the literature [17, 42, 43, 44, 45, 98]. However, the specific structure of the mathematical approach, the usage of just a known phenotype distribution, and the connection to the effectiveness of targeted recruitment strategies (see Section 3.2.4) are new.

3.2.2   Motivational example

Figure 3.1 shows a probability diagram describing the relation between the RhD genotype of a donor and its parents and siblings (i.e., brothers or sisters). The probability that a donor has a particular genotype is the a priori probability. From the figure it is clear that the probability of a sibling having the same genotype requires information on genotypes of the parents. However, it might be that the distribution of genotypes in the general population is unknown or difficult to obtain. On the other hand, the phenotype distribution for the general population is usually more easily available, so it would be convenient if we could use this instead, to determine the genotype distribution. This is possible by using quadratic stochastic operators.

When the a priori probabilities are known, Bayes rule is applied to find the probability that a relative of a donor has a specific RhD genotype, given the genotype of the donor. In order to compute these probabilities, particularly for a sibling of a donor, we thus need to work top-down.

3.2.3   Finding a stationary distribution

For a particular blood group, a child inherits its genotype from a combination of genotypes of the parents. For the RhD blood group a genotype consists of two genes, each of which either d or D, leading to three possible genotype combinations: dd, Dd, and DD. Each parent gives one of these two to the child. The probability that two parents with a particular genotype conceive a child with a certain genotype is captured by an inheritance matrix P. For the RhD blood groups, the inheritance matrix is depicted in Figure 3.2. We use this inheritance

Figure 3.1: Probability diagram which relates the RhD genotype of a donor to the RhD genotypes of donor's parents and siblings
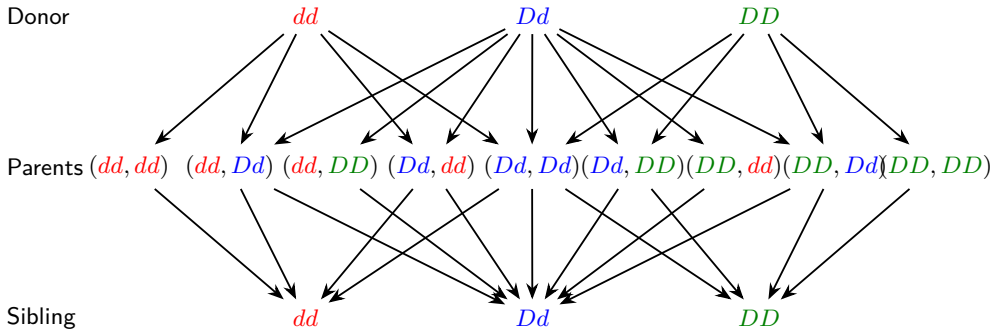
Figure 3.2: Inheritance matrix P for the RhD blood group.
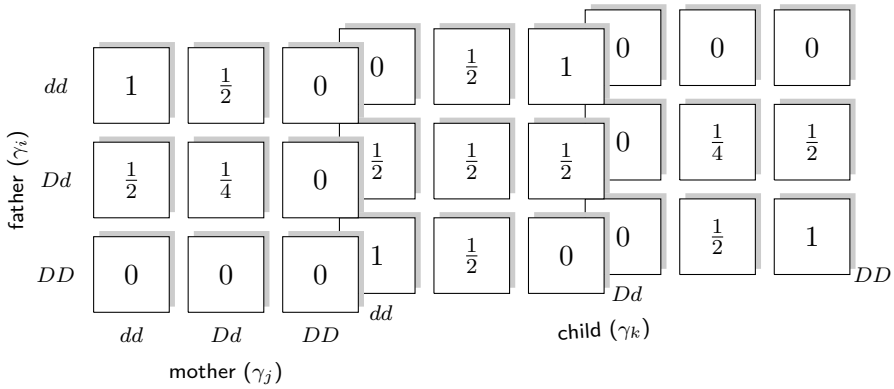
matrix $P \in \mathbb{R}^{3 \times 3 \times 3}$ to compute a stationary distribution of genotypes. The exact structure of this matrix is explained later in Section 3.3.1.

Let $\boldsymbol{x}^{(n)} \in \mathbb{R}^{3 \times 1}$ be a column vector containing the genotype distribution for the RhD blood groups in generation n. We assume that this genotype distribution is stationary, which implies that the distribution of genotypes in generation $n-1$ is equal to the distribution of genotypes in generation n: $\boldsymbol{x}^{(n-1)} = \boldsymbol{x}^{(n)} = \boldsymbol{x}$. Let $\boldsymbol{x}_{\text{father}}$, $\boldsymbol{x}_{\text{mother}}$, and $\boldsymbol{x}_{\text{child}}$ be the genotype distributions of respectively father, mother, and child. Then, in a stationary population, the following equation holds:

$$\boldsymbol{x}_{\text{father}}^{\top} P \boldsymbol{x}_{\text{mother}} = \boldsymbol{x}_{\text{child}} \quad \Rightarrow \quad \boldsymbol{x}^{\top} P \boldsymbol{x} = \boldsymbol{x}. \tag{3.2}$$

Moreover, the genotypes are related to the phenotypes. This relation was given in the matrix S (equation (3.1)). Besides equation (3.2) the following equation

should also hold for variable $\boldsymbol{x}$:

$$S^\top \boldsymbol{x} = \boldsymbol{f}, \tag{3.3}$$

where $\boldsymbol{f} \in \mathbb{R}^{2\times 1}$ is the phenotype distribution. For the RhD blood groups, equations (3.2) and (3.3) can be solved analytically, which gives:

$$\begin{cases} x_{dd}^2 + x_{dd}x_{Dd} + \frac{1}{4}x_{Dd}^2 & = x_{dd} \\ x_{dd}x_{Dd} + 2x_{dd}x_{DD} + \frac{1}{2}x_{Dd}^2 + x_{Dd}x_{DD} = x_{Dd} \\ \frac{1}{4}x_{Dd}^2 + x_{Dd}x_{DD} + x_{DD}^2 & = x_{DD} \\ x_{dd} & = f_d \\ x_{Dd} + x_{DD} & = f_D \end{cases} \Rightarrow \begin{cases} x_{dd} & = f_d \\ x_{Dd} & = f_D - \left(1 - \sqrt{f_d}\right)^2 \\ x_{DD} & = \left(1 - \sqrt{f_d}\right)^2 \end{cases}.$$

Note that this analytic solution is in accordance with the Hardy-Weinberg law [53]. Since $\boldsymbol{f}^\top = (f_d, f_D) = (0.168, 0.832)$ we get

$$\boldsymbol{x} = \begin{bmatrix} x_{dd} \\ x_{Dd} \\ x_{DD} \end{bmatrix} = \begin{bmatrix} 0.168 \\ 0.484 \\ 0.348 \end{bmatrix}. \tag{3.4}$$

In a similar way equations and computations can be provided for the ABO blood group system from which we find

$$\boldsymbol{x} = \begin{bmatrix} x_{OO} \\ x_{OA} \\ x_{OB} \\ x_{AA} \\ x_{AB} \\ x_{BB} \end{bmatrix} = \begin{bmatrix} 0.440 \\ 0.358 \\ 0.087 \\ 0.073 \\ 0.038 \\ 0.004 \end{bmatrix}. \tag{3.5}$$

In Casas et al. [17] square root expressions have been provided for the ABO blood group system and are therefore omitted here. However, this reference has not discussed the concept of effectiveness. This will be elaborated on in the next section.

3.2.4   Effectiveness of recruiting next of kin for donorship

Donors are recruited for their phenotypes expressions (blood is matched on phenotypes), however inheritance is determined by genotypes. Therefore, to compute the probability that a sibling of a donor with a particular phenotype has the same phenotype, the stationary genotype distribution is required. Once this genotype distribution has been obtained the likelihood of a particular blood group for a

sibling, given the blood group of a relative, can then be computed using Bayes' rule.

Suppose that we have a RhD-pos donor, the likelihood that its sibling is also RhD-pos can be computed by calculation the conditional probability:

$$\mathbb{P}\left(\text{sibling D} \mid \text{donor D}\right) = 0.898. \tag{3.6}$$

We find that the conditional probability is slightly higher than the probability that a random individual is RhD-pos (0.850). The effectiveness, defined as the difference between these two probabilities, is equal to

$$\begin{aligned} E_D &= \mathbb{P}\left(\text{sibling D} \mid \text{donor D}\right) - f_D \\ &= 0.066. \end{aligned} \tag{3.7}$$

Figure 3.3 shows the results of an analysis of ABO and RhD blood groups. Especially for rare blood groups (i.e. B, AB and RhD-neg) it appears to be effective to recruit among relatives. Here, the likelihood of a similar blood group is considerably higher than that of the general population. For example, for the RhD-neg blood group the likelihood increases from 0.168 to 0.410 for parents and to 0.497 for siblings. Note that the probability of the siblings is higher than that of the parents.

The most important ABO, RhD blood group is O, RhD-neg, since this is the blood group of a so-called universal donor. This means that every individual can receive RBC's from a donor with this blood group. Figure 3.3 shows that recruiting O, RhD-neg donors among relatives of donors with an O, RhD-neg blood group is five or four times more effective for siblings and parents respectively, than recruiting donors at random. These computations are insightful when assessing targeted donor recruitment among relatives.

This section provided an illustration of calculating next of kin blood group probabilities for the ABO, RhD blood groups. In the next sections we will provide a more generic mathematical framework to compute i) stationary genotype distribution and ii) effectiveness of recruiting next of kin for blood donorship by using quadratic stochastic operators and Bayesian statistics. This allows calculating next of kin probabilities for more complex blood group combinations.

3.3   Generic mathematical approach

As illustrated in Section 3.2.4 we want to calculate the conditional probability that a relative of a donor has the same phenotype as the donor, or mathematically stated:

$$\mathbb{P}\left(\text{relative } \varphi \mid \text{donor } \varphi\right), \tag{3.8}$$

Figure 3.3: Proportion of ABO, RhD blood groups in the general population and conditional probabilities that parents/siblings have the same blood group. The numbers above the conditional probabilities represent the effectiveness of recruiting relatives of a donor with a known blood group, where the effectiveness is defined as the difference between the the proportion of individuals with a particular blood group in the population (see Table 3.1) and the conditional probabilities.
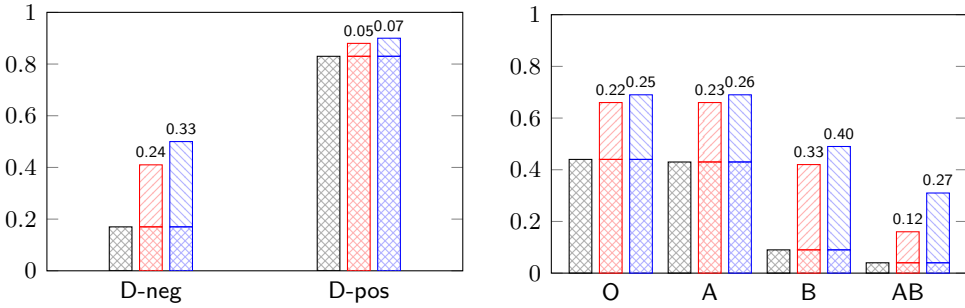
(a) Proportion and inheritance probability of RhD blood groups

(b) Proportion and inheritance probability of ABO blood groups
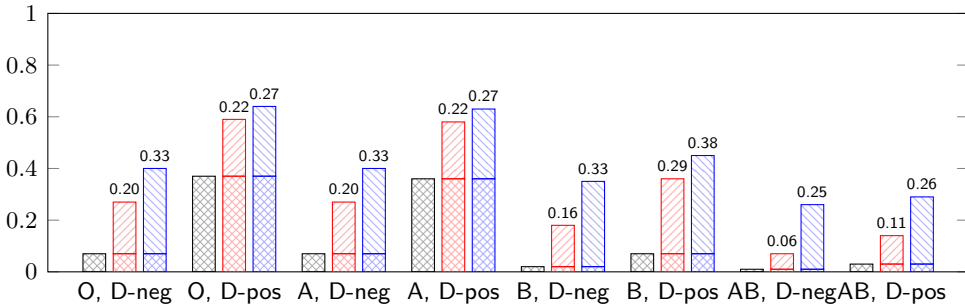
(c) Proportion and inheritance probability of ABO, D blood groups

where $\varphi \in \mathcal{F}$ is the known phenotype of the donor. We therefore aim to provide a unified mathematical framework, starting in this section with providing a generic mathematical approach for computing the stationary genotype distribution. In the next section it will be shown how this stationary genotype distribution is used to calculated the effectiveness of recruiting next of kin for blood donorship.

First, in Section 3.1.1, we start with mathematically modelling the blood group genetics and introduce some notation. Next, In Section 3.3.2 the calculation steps required are listed and in the remainder of Section 3.3 we elaborate on the computation of a stationary genotype distribution.

3.3.1 Blood group genetics

To explain the relation between the blood group of a child and its parents we start with a compact description of the underlying genetic mechanism of inheritance. The information of an individual's blood group is present on the genes, which occur in pairs on homologous chromosomes at particular positions called loci (singular: locus). Genes that occur at the same locus are allelic to each other and are therefore also referred to as alleles. Each allele may encode for the production of a specific antigen. For example, the ABO blood groups are determined by three alleles A, B, and O, where A encodes for the production of antigen A, B encodes for the production of antigen B, and O encodes for no antigen production. To write this down mathematically we first introduce for each locus a set of alleles $\mathcal{L}$ and a set of antigens $\mathcal{A}$. Then, for each allele $\ell \in \mathcal{L}$ a binary vector of length $|\mathcal{A}|$ is constructed with $\ell_a = 1$ if allele ℓ encodes for the production of antigen $a \in \mathcal{A}$ and $\ell_a = 0$ otherwise. Finally, the alleles are sorted into colexicographical order (expressed as '$\preceq$'), i.e., $(\mathcal{L}_{\mathrm{ABO}}, \preceq) = \{O, A, B\} = \{[0\ 0], [1\ 0], [0\ 1]\}$. Hence, $(\mathcal{L}, \preceq)$ is a colexicographically ordered set of alleles.

In contrast to the ABO blood groups, which are determined by alleles lying on a single locus, the Rhesus blood groups are determined by a combination of alleles occurring at multiple loci. This combination of alleles is called a haplotype (the multilocus analogue of an allele at a single locus), where a haplotype consists of one allele from each of the loci. The set of haplotypes is denoted by $\mathcal{H}$, where each $h \in \mathcal{H}$ can be written as a union of alleles belonging to unique loci. For example, the set of haplotypes for the Rhesus blood groups is determined by alleles from three loci ($\mathcal{L}_{\mathrm{D}} = \{D, d\} = \{[1], [0]\}$, $\mathcal{L}_{\mathrm{C}} = \{C, c\} = \{[1\ 0], [0\ 1]\}$, and $\mathcal{L}_{\mathrm{E}} = \{E, e\} = \{[1\ 0], [0\ 1]\}$) leading to eight different Rhesus haplotypes:

binary representation	antigens	haplotype
$[0\ 0\ 1\ 0\ 1]$	ce	dce
$[0\ 0\ 1\ 1\ 0]$	cE	dcE
$[0\ 1\ 0\ 0\ 1]$	Ce	dCe
$[0\ 1\ 0\ 1\ 0]$	CE	dCE
$[1\ 0\ 1\ 0\ 1]$	Dce	Dce
$[1\ 0\ 1\ 1\ 0]$	DcE	DcE
$[1\ 1\ 0\ 0\ 1]$	DCe	DCe
$[1\ 1\ 0\ 1\ 0]$	DCE	DCE

Although the sets $\mathcal{L}_{\mathrm{D}}$, $\mathcal{L}_{\mathrm{C}}$, and $\mathcal{L}_{\mathrm{E}}$ all consist of two alleles, they are different. On the one hand, the alleles in the sets $\mathcal{L}_{\mathrm{C}}$ and $\mathcal{L}_{\mathrm{E}}$ always lead to the production of antigens, i.e. $[1\ 0] \in \mathcal{L}_{\mathrm{C}}$ implies production of antigens C, $[0\ 1] \in \mathcal{L}_{\mathrm{C}}$ implies production of antigens c, $[1\ 0] \in \mathcal{L}_{\mathrm{E}}$ implies production of antigens E, and $[0\ 1] \in \mathcal{L}_{\mathrm{E}}$ implies production of antigens e. On the other hand, the alleles in the set $\mathcal{L}_{\mathrm{D}}$

might lead to the production of an antigen, i.e. $[1] \in \mathcal{L}_D$ implies production of antigens D, but $[0] \in \mathcal{L}_D$ implies that no antigens are produced. The set $\mathcal{H}$ has cardinality $|\mathcal{H}| = \prod_i |\mathcal{L}_i|$.

Define $\mathcal{G} = \{\gamma_1, ..., \gamma_m\}$ as a set of genotypes consisting of all combinations of 2 haplotypes from $\mathcal{H}$:

$$\mathcal{G} = \{ \underbrace{\{h_1, h_1\}, ..., \{h_1, h_{|\mathcal{H}|}\}}_{|\mathcal{H}| \text{ elements}}, \underbrace{\{h_2, h_2\}, ..., \{h_2, h_{|\mathcal{H}|}\}}_{|\mathcal{H}|-1 \text{ elements}}, ..., \underbrace{\{h_{|\mathcal{H}|}, h_{|\mathcal{H}|}\}}_{1 \text{ element}}\}, \quad (3.9)$$

with cardinality $m = |\mathcal{G}| = \frac{1}{2}|\mathcal{H}|(|\mathcal{H}| + 1)$. Finally, let $\mathcal{F} = \{\varphi_1, ..., \varphi_n\}$ be the lexicographic ordered set of phenotypes, with cardinality $n = |\mathcal{F}|$. These phenotypes determine which antigens are present on the red blood cells. Define $S \in \{0, 1\}^{m \times n}$ as a matrix describing the relation between genotypes and phenotypes, that is

$$S_{ij} = \begin{cases} 1, & \text{if genotype } \gamma_i \in \mathcal{G} \text{ leads to phenotype } \varphi_j \in \mathcal{F}, \\ 0, & \text{otherwise.} \end{cases} \quad (3.10)$$

Children inherit blood group antigens from their parents. Which antigens are inherited depends on the genotypes of both parents. Suppose that the father has genotype $\gamma_i \in \mathcal{G}$ ($\gamma_i = \{h_{i_1}, h_{i_2}\}$), the mother has genotype $\gamma_j \in \mathcal{G}$ ($\gamma_j = \{h_{j_1}, h_{j_2}\}$) and they have a child with genotype $\gamma_k \in \mathcal{G}$. Clearly, this child could have four different genotypes, since there are four different combinations of the haplotypes of the parents: $\{h_{i_1}, h_{j_1}\}$, $\{h_{i_1}, h_{j_2}\}$, $\{h_{i_2}, h_{j_1}\}$, and $\{h_{i_2}, h_{j_2}\}$. We assume that Mendelian rules hold, which implies that each combination occurs with probability $\frac{1}{4}$. In this section, we will index the the genotypes by

- γ_i - genotype of the father,

- γ_j - genotype of the mother,

- γ_k - genotype of the child,

and use no index if we do not refer specifically to a father, mother, or child.

In order to construct the inheritance matrix $P \in \mathbb{R}^{m \times m \times m}$ we first introduce vectors $\boldsymbol{v}_h = [v_h(\gamma_1), ..., v_h(\gamma_m)]$, $h \in \mathcal{H}$, where $v_h(\gamma_i)$ is the probability that a parent with genotype $\gamma_i = \{h_{i_1}, h_{i_2}\}$ will give haplotype h to the child:

$$v_h(\gamma_i) = \begin{cases} 1, & \text{if } \gamma_i = \{h, h\}, \\ \frac{1}{2}, & \text{if } \gamma_i = \{h, \not h\} \text{ or } \gamma_i = \{\not h, h\}, \\ 0, & \text{otherwise.} \end{cases} \quad h \in \mathcal{H}, \quad \gamma_i \in \mathcal{G}, \quad (3.11)$$

Then the probability that a child has genotype $\gamma_k = \{h_i, h_j\}$, where h_i is the haplotype the child inherited from the father and h_j is the haplotype the child

Figure 3.4: Inheritance matrix P, with $P\left(\gamma_k \mid \gamma_i, \gamma_j\right)$ the probability that two individuals with genotypes $\gamma_i \in \mathcal{G}$ and $\gamma_j \in \mathcal{G}$ conceive a child with genotype $\gamma_k \in \mathcal{G}$.
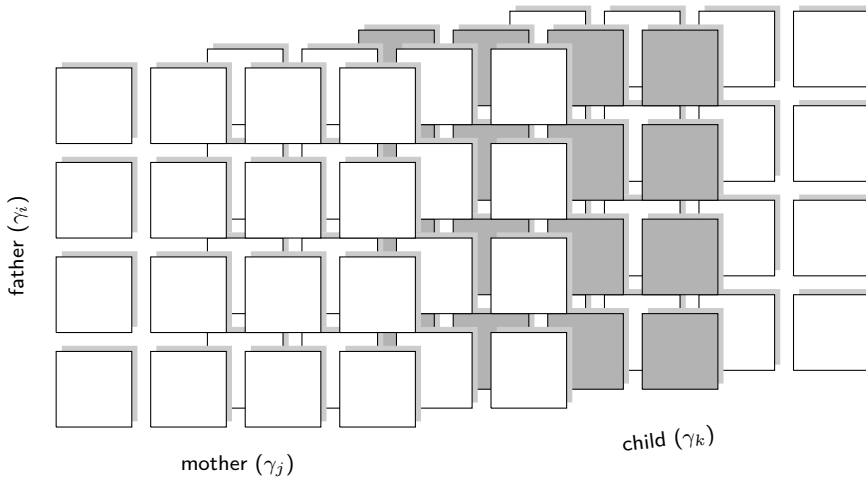
inherited from the mother, is equal to:

$$P\left(\gamma_k\right) = P\left(\{h_i, h_j\}\right) = \begin{cases} \boldsymbol{v}_{h_i}\boldsymbol{v}_{h_j}^T, & \text{if } i = j, \\ \boldsymbol{v}_{h_i}\boldsymbol{v}_{h_j}^T + \boldsymbol{v}_{h_j}\boldsymbol{v}_{h_i}^T, & \text{if } i \neq j. \end{cases} \tag{3.12}$$

Note that $P(\gamma_k) \in \mathbb{R}^{m \times m}$ is a two dimensional matrix as is shown in Figure 3.4.

3.3.2 Steps

The probability that two relatives share the same blood group is substantially higher than the probability that two individuals from the general population share the same blood group. For selective donor recruitment it is therefore worthwhile to quantify these probabilities as a function of the family relation. One might have the perception that these probabilities can be easily computed by elementary statistics. This is true, except that the a priori probabilities, i.e. the genotype distributions, are generally unknown and have to be calculated first. As will be shown, these a priori probabilities can be determined by a system of quadratic equations or rather a system of quadratic stochastic operators. Therefore, the mathematical approach, combining both elementary statistics and operations research related methods, can be divided into the following three steps:

- Determine the stationary distribution of genotypes.

- Compute the probability that a relative of a donor has a particular phenotype given that this donor has a particular phenotype.

- Compute the effectiveness of recruiting a next of kin donor instead of an individual from the general population.

3.3.3   Determine a stationary distribution of genotypes

By performing simple tests it is possible to determine the distribution of phenotypes in a population for a (combination of) blood group(s). Genotype distributions or allele frequencies are more difficult to obtain. A way to obtain estimates of these genotype distributions is by using QSO's. These estimates are based on the known phenotype distributions and the assumption that the genotype distributions within a population are stable.

First, we explain how we can model the inheritance of antigens by using a quadratic stochastic operator. This leads to a system of quadratic equations. Next, we show how this system of quadratic equations can be solved, by iteratively solving a least squares problem.

Consider a set $\mathcal{G}$ of genotypes. Let x_γ be a variable that describes the frequency of genotype $\gamma \in \mathcal{G}$ in a population and let $P\left(\gamma_k \mid \gamma_i, \gamma_j\right)$ be the probability that two individuals with genotypes $\gamma_i \in \mathcal{G}$ and $\gamma_j \in \mathcal{G}$ conceive a child with genotype $\gamma_k \in \mathcal{G}$. Now, as in Section 3.2.3, let $x = x_{\text{father}} = x_{\text{mother}} = x_{\text{child}}$, then the following equations hold:

$$x_{\text{father}}^\top P x_{\text{mother}} = x_{\text{child}} \quad \Rightarrow \quad x^\top P x = x, \tag{3.13}$$

where $P\left(\gamma_k \mid \gamma_i, \gamma_j\right)$ is the heredity matrix which satisfies the following three properties: $P\left(\gamma_k \mid \gamma_i, \gamma_j\right) \geq 0$, $P\left(\gamma_k \mid \gamma_i, \gamma_j\right) = P\left(\gamma_k \mid \gamma_j, \gamma_i\right)$, $\sum_{\gamma_k \in \mathcal{G}} P\left(\gamma_k \mid \gamma_i, \gamma_j\right) = 1$.

Since we have a system of quadratic equations, there could be multiple stationary solutions x. Based on the phenotype distribution f we can investigate which of these solutions is correct, requiring $S^\top x = f$. Hence, we need to solve the following system of equations:

$$\begin{cases} x^\top P x = x, \\ S^\top x = f. \end{cases} \tag{3.14}$$

To compute a solution x that satisfies (3.14) we are first going to rewrite this system of quadratic equations as:

$$\begin{cases} x^\top P x = x \\ S^\top x = f \end{cases} \quad \Rightarrow \quad \begin{cases} \left(x^\top P - I\right) x = 0 \\ S^\top x = f \end{cases} \quad \Rightarrow \quad \underbrace{\begin{bmatrix} x^\top P - I \\ S^\top \end{bmatrix}}_{A(x)} x = \underbrace{\begin{bmatrix} 0 \\ f \end{bmatrix}}_{b}, \tag{3.15}$$

where $I \in \mathbb{R}^{m \times m}$ is the identity matrix and $\mathbf{0} \in \mathbb{R}^{m \times 1}$ is the zero vector. In short, we thus get

$$A(\boldsymbol{x})\boldsymbol{x} = \boldsymbol{b}, \tag{3.16}$$

with $A(x) \in \mathbb{R}^{(m+n) \times m}$, $\boldsymbol{x} \in \mathbb{R}^{m \times 1}$, and $\boldsymbol{b} \in \mathbb{R}^{(m+n) \times 1}$. Since this is not a linear but an implicit equation, an iterative approximate procedure is proposed. For given $\boldsymbol{x}_0$ let $\boldsymbol{x}^{(n)}$ for $n = 1, 2, ...$ be determined by

$$A(\boldsymbol{x}^{(n-1)})\boldsymbol{x}^{(n)} = b. \tag{3.17}$$

In Section 3.3.5 we will make this more explicit. To this end, since there are different methods possible, let us first provide the one that will be used.

3.3.4   QR-factorization

To support iteration (3.17), if we regard the matrix A independent of $\boldsymbol{x}^{(n-1)}$, then we just have a system of linear equations. Normally, a linear system can easily be solved by applying Gaussian elimination. However, in this specific case an exact solution may not exist, since this system has more equations than unknown variables. Accordingly, we propose to solve a least squares problem. Let $\boldsymbol{r} = A\boldsymbol{x}^{(n)} - \boldsymbol{b}$, or simply $\boldsymbol{r} = A\boldsymbol{x} - \boldsymbol{b}$, be the vector of residuals. Next, we want to find a solution $\boldsymbol{x}$ that minimizes the sum of squared residuals:

$$\min_{\boldsymbol{x} \in \mathbb{R}^m} \left\{ \|\boldsymbol{r}\|_2^2 \mid \boldsymbol{r} = A\boldsymbol{x} - \boldsymbol{b} \right\}. \tag{3.18}$$

Different methods are known to solve least squares problems. One of them, based upon QR factorization [114], will be applied here. If the matrix A has full column rank, then it can be decomposed into the matrices Q and R ($A = QR$), such that the matrix $Q \in \mathbb{R}^{(m+n) \times (m+n)}$ has orthonormal columns and the matrix $R \in \mathbb{R}^{(m+n) \times m}$ is upper triangular (see Figure 3.5). In Section 3.8 (Appendix II) we prove that the matrix A has indeed full column rank and hence the residuals can be written as $\boldsymbol{r} = QR\boldsymbol{x} - \boldsymbol{b}$.

Define $\bar{\boldsymbol{r}} = Q^\top \boldsymbol{r}$ as a linear transformation of the residuals. Then minimizing $\|\bar{\boldsymbol{r}}\|_2^2$ is equivalent to minimizing $\|\boldsymbol{r}\|_2^2$, since

$$\|\bar{\boldsymbol{r}}\|_2^2 = \left(Q^\top \boldsymbol{r} \right)^\top Q^\top \boldsymbol{r} = \boldsymbol{r}^\top Q Q^\top \boldsymbol{r} = \boldsymbol{r}^\top \boldsymbol{r} = \|\boldsymbol{r}\|_2^2. \tag{3.19}$$

Moreover, note that $\bar{\boldsymbol{r}} = R\boldsymbol{x} - Q^\top \boldsymbol{b}$ and hence (3.18) is equivalent to

$$\min_{\boldsymbol{x} \in \mathbb{R}^m} \left\{ \|\bar{\boldsymbol{r}}\|_2^2 \mid \bar{\boldsymbol{r}} = R\boldsymbol{x} - Q^\top \boldsymbol{b} \right\}. \tag{3.20}$$

We can find an exact solution to (3.20) by exploiting the specific structure of R. Partition R into an upper triangular matrix R_1 and a zero matrix. Similarly, we can

Figure 3.5: QR decomposition of the matrix A, where Q is a orthogonal matrix and R is an upper triangular matrix. Matrices Q and R can be partitioned, such that $Q = [Q_1 \ Q_2]$, $R^\top = [R_1 \ 0]$, and R_1 is an upper triangular matrix.
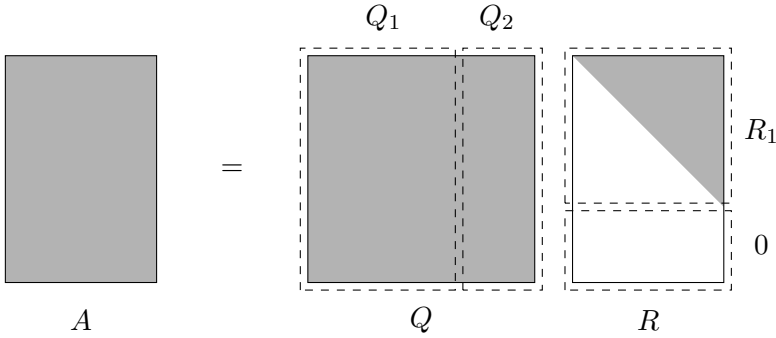
write $Q = [Q_1 \ Q_2]$ and $\bar{r}^\top = \begin{bmatrix} \bar{r}_1^\top & \bar{r}_2^\top \end{bmatrix}$ (see Figure 3.5). Hence, $\bar{r} = Rx - Q^\top b$ can be split into two sets of equations

$$\begin{cases} \bar{r}_1 = R_1 x - Q_1^\top b, \\ \bar{r}_2 = -Q_2^\top b, \end{cases} \tag{3.21}$$

and (3.20) is equivalent to

$$\left\| -Q_2^\top b \right\|_2^2 + \min_{x \in \mathbb{R}^m} \left\{ \|\bar{r}_1\|_2^2 \mid \bar{r}_1 = R_1 x - Q_1^\top b \right\}. \tag{3.22}$$

Note that $Q_1^\top b = R_1 x$ consist of m equations with m unknowns and can be solved by backward substitution since R_1 is an upper triangular matrix. Or by multiplying both side by R_1^{-1}: $x = R_1^{-1} Q_1^\top b$. Hence, the second part of (3.22) is equal to zero and therefore the sum of squared residuals is equal to

$$\|\bar{r}\|_2^2 = \left\| Q_2^\top b \right\|_2^2. \tag{3.23}$$

This implies that if $\left\| Q_2^\top b \right\|_2^2 = 0$ all equations $Ax = b$ are satisfied. Moreover the minimizer of (3.20) is equal to

$$x = R_1^{-1} Q_1^\top b. \tag{3.24}$$

3.3.5 Iterative procedure

In Section 3.3.4 we included $x^{(n-1)}$ in A. This suggests the following iterative procedure: choose an initial solution $x^{(0)}$ and find a new solution $x^{(1)}$ by solving the least square problem described in the previous section. Hence, a solution to equations (3.17) can be found by iteratively solving

$$x^{(n)} = R_1 \left(x^{(n-1)} \right)^{-1} Q_1 \left(x^{(n-1)} \right)^{\top} b, \quad n = 1, 2, ..., \tag{3.25}$$

where the sum of squared residuals equals

$$\|r\|_2^2 = \left\| Q_2 \left(x^{(n-1)} \right)^{\top} b \right\|_2^2, \quad n = 1, 2, ... \tag{3.26}$$

A solution $x^{(n)}$ is defined satisfactory when $\left\| x^{(n)} - x^{(n-1)} \right\|_1 < 10^{-6}$. Hence, we consider (3.16) to be solved numerically.

3.3.6 Convergence of the algorithm

To support the iterative procedure from Section 3.3.6 we have performed numerical experiments for different blood group systems (e.g., ABO, Rhesus, Kell, Duffy, Kidd, MNS), phenotype distributions, and populations (e.g., European, African). We took $x^{(0)} \in X_0$, where X_0 is the set of all identity vectors of size $|\mathcal{G}|$. This means that every starting position $x^{(0)}$ represents an initial population with only one genotype. Despite these extreme starting points all experiments converged to the same stationary solution x within 37 iterations. For the Rhesus system, which will be presented in Section 3.5.1 there are 36 starting points. In Figure 3.6, the fasted, slowest, and average convergence rates from these experiments for the Rhesus system are shown.

3.4 Effectiveness of recruiting next of kin for blood donorship

The effectiveness of recruiting a next of kin for donorship for phenotype $\varphi \in \mathcal{F}$ is defined as the difference between the conditional probability that a relative of a donor has the same phenotype φ as the donor and the probability that an arbitrary individual in the population has this phenotype, i.e.,

$$E_\varphi = \mathbb{P}\left(\text{relative } \varphi \mid \text{donor } \varphi\right) - \mathbb{P}\left(\text{individual } \varphi\right), \qquad \forall \varphi \in \mathcal{F}. \tag{3.27}$$

We would like to rewrite this equation into the known phenotype distribution (f), the heredity matrix (P), and the stationary genotype distribution (x).

Figure 3.6: Convergence speed of the iterative procedure for the Rhesus blood group system.
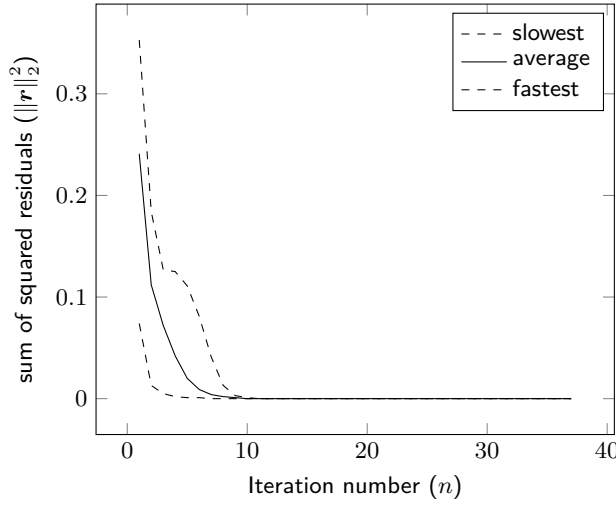
The conditional probability in equation (3.27) is, according to Bayes' rule, equal to

$$
\begin{aligned}
\mathbb{P}\left(\text{relative } \varphi \mid \text{donor } \varphi\right) &= \frac{\mathbb{P}\left(\text{donor } \varphi \,\cap\, \text{relative } \varphi\right)}{\mathbb{P}\left(\text{donor } \varphi\right)}, \\
&= \tfrac{1}{f_\varphi}\mathbb{P}\left(\text{donor } \varphi \,\cap\, \text{relative } \varphi\right).
\end{aligned}
\tag{3.28}
$$

Since every genotype is related to a single phenotype the probability that the donor and its relative have the same phenotype can be computed by summing over all combinations of genotypes that they can have. Hence, if we denote the genotype of the donor by $\gamma_i \in \mathcal{G}$ and the genotype of its relative by $\gamma_j \in \mathcal{G}$, then we should sum over those combinations of genotypes for which $S_{i\varphi}$ and $S_{j\varphi}$ are both equal to one:

$$
\mathbb{P}\left(\text{relative } \varphi \mid \text{donor } \varphi\right) = \frac{1}{f_\varphi} \sum_{\substack{\gamma_i \in \mathcal{G} \\ S_{i\varphi}=1}} \sum_{\substack{\gamma_j \in \mathcal{G} \\ S_{j\varphi}=1}} \mathbb{P}\left(\text{relative } \gamma_j \,\cap\, \text{donor } \gamma_i\right).
\tag{3.29}
$$

Applying Bayes' rule for the second time gives

$$
\begin{aligned}
\mathbb{P}\left(\text{relative } \varphi \mid \text{donor } \varphi\right) &= \tfrac{1}{f_\varphi} \sum_{\substack{\gamma_i \in \mathcal{G} \\ S_{i\varphi}=1}} \sum_{\substack{\gamma_j \in \mathcal{G} \\ S_{j\varphi}=1}} \left[\mathbb{P}\left(\text{donor } \gamma_i\right) \cdot \mathbb{P}\left(\text{relative } \gamma_j \mid \text{donor } \gamma_i\right)\right], \\
&= \tfrac{1}{f_\varphi} \sum_{\substack{\gamma_i \in \mathcal{G} \\ S_{i\varphi}=1}} \sum_{\substack{\gamma_j \in \mathcal{G} \\ S_{j\varphi}=1}} \left[x_{\gamma_i} \cdot \mathbb{P}\left(\text{relative } \gamma_j \mid \text{donor } \gamma_i\right)\right],
\end{aligned}
$$

Relative	$\mathbb{P}\left(\text{relative } \gamma_j \mid \text{donor } \gamma_i\right)$
Sibling	$\displaystyle\sum_{\gamma_v \in \mathcal{G}} \sum_{\gamma_w \in \mathcal{G}} P\left(\gamma_j \mid \gamma_v, \gamma_w\right) \cdot \frac{x_{\gamma_v} x_{\gamma_w} P\left(\gamma_i \mid \gamma_v, \gamma_w\right)}{x_{\gamma_i}}$
Parent	$\displaystyle\sum_{\gamma_v \in \mathcal{G}} \frac{x_{\gamma_j} x_{\gamma_v} P\left(\gamma_i \mid \gamma_j, \gamma_v\right)}{x_{\gamma_i}}$
Child	$\displaystyle\sum_{\gamma_v \in \mathcal{G}} x_{\gamma_v} P\left(\gamma_i \mid \gamma_j, \gamma_v\right)$
Uncle / Aunt	$\displaystyle\sum_{\gamma_v \in \mathcal{G}} \left[\mathbb{P}\left(\text{sibling } \gamma_j \mid \text{parent } \gamma_v\right) \cdot \mathbb{P}\left(\text{parent } \gamma_v \mid \text{donor } \gamma_i\right)\right]$
Nephew / Niece	$\displaystyle\sum_{\gamma_v \in \mathcal{G}} \left[\mathbb{P}\left(\text{child } \gamma_j \mid \text{sibling } \gamma_v\right) \cdot \mathbb{P}\left(\text{sibling } \gamma_v \mid \text{donor } \gamma_i\right)\right]$
Grandparent	$\displaystyle\sum_{\gamma_v \in \mathcal{G}} \left[\mathbb{P}\left(\text{parent } \gamma_j \mid \text{parent } \gamma_v\right) \cdot \mathbb{P}\left(\text{parent } \gamma_v \mid \text{donor } \gamma_i\right)\right]$

Table 3.2: Probability that a relative of a donor has genotype γ_j given that the donor has genotype γ_i.

$$\tag{3.30}$$

where $\mathbb{P}\left[\text{relative } \gamma_j \mid \text{donor } \gamma_i\right]$ can be expressed in terms of P and x according to the relation between the donor and its relative, as is indicated in Table 3.2. The second probability in equation (3.27) equals the frequency of φ in the general population. Hence, the effectiveness of recruiting a next of kin for donorship for phenotype $\varphi \in \mathcal{F}$ is equal to

$$E_\varphi = \frac{1}{f_\varphi} \sum_{\substack{\gamma_i \in \mathcal{G} \\ S_{i\varphi}=1}} \sum_{\substack{\gamma_j \in \mathcal{G} \\ S_{j\varphi}=1}} \left[x_{\gamma_i} \cdot \mathbb{P}\left(\text{relative } \gamma_j \mid \text{donor } \gamma_i\right)\right] - f_\varphi. \tag{3.31}$$

3.5 Application to multiple blood groups

In Section 3.2 we illustrated the effectiveness of a targeted donor recruitment strategy for siblings and parents of donors with particular ABO, D blood groups. To demonstrate the generic feature of our mathematical approach we analyse the more complicated Rhesus blood group system. Patients with sickle cell disease or thalassemia require regular (life-long) blood transfusions. To prevent these

Relative	Effectiveness		
	minimum	average	maximum
Sibling	0.197	0.275	0.361
Parent	0.068	0.156	0.245
Child	0.068	0.156	0.245
Uncle / Aunt	0.028	0.078	0.122
Nephew / Niece	0.028	0.078	0.122
Grandparent	0.034	0.078	0.122

Table 3.3: Effectiveness of recruiting new donors among relatives of donors with a known Rh phenotype. Whereas this effectiveness depends on the specific phenotype sought for we reported a minimum, average and maximum effectiveness measure.

recipients from forming antibodies against foreign RBC antigens, they are matched for a relatively large number of antigens. However, it is not easy to ensure that there is a sufficient number of required blood units available. We show how our generic model can be used to find more donors with the desired blood group combinations.

3.5.1   Rh blood group system

In contrast to the well known RhD blood group consisting of just 3 genotypes and 2 phenotypes, the full Rh system consists of 36 different genotypes and 18 different phenotypes. To compute the effectiveness of recruiting relatives of donors with desired Rh phenotypes, we first have to computed the stationary distribution based on Rh phenotype probabilities.

We computed these stationary probabilities by solving equations (3.14) via an iterative procedure (3.25) (see Table 3.4). Next, we can apply use Bayes rule to compute the effectiveness of recruiting relatives as compared to random individuals from the general population. In Table 3.3 the calculated effectiveness of recruiting new donors among relatives is presented. Dependent of the family relationship effectiveness changes. Note that the effectiveness is variable, as it is dependent on the phenotype considered.

3.5.2   Recruitment of special blood groups

In the Netherlands, for some patient groups (e.g. women of reproductive age, patient with hemoglobinopathies) blood for transfusion is matched for up to thirteen antigens. With current recruitment strategies it can be difficult to find enough donors with particular blood groups combinations. Moreover, due to the fact that the donor population in the Netherlands is mainly Caucasian, the patient population increasingly diversifying, and blood group frequencies differ be-

Phenotype	f	Genotype	x
DCce	0.349	dce /DCe	0.326
		Dce /DCe	0.022
		dCe /Dce	0.001
DCe	0.185	DCe/DCe	0.176
		dCe /DCe	0.009
ce	0.151	dce /dce	0.151
DCcEe	0.133	DcE/DCe	0.119
		dcE /DCe	0.010
		dCe /DcE	0.003
		dce /DCE	0.002
		dCE/Dce	0.000
		Dce /DCE	0.000
DcEe	0.118	dce /DcE	0.110
		Dce /DcE	0.007
		dcE /Dce	0.001
DcE	0.023	DcE/DcE	0.020
		dcE /DcE	0.003
Dce	0.021	dce /Dce	0.020
		Dce /Dce	0.001
cEe	0.009	dcE /dce	0.009
Cce	0.008	dCe /dce	0.008
DCEe	0.002	DCe/DCE	0.002
		dCe /DCE	0.000
DCcE	0.001	DcE/DCE	0.001
		DcE/dCE	0.000
other	0.000		

Table 3.4: Phenotype distribution of the Rh blood group system and the corresponding genotype distribution computed by the generic mathematical approach.

tween ethnic populations, the likelihood of finding suitable blood units for non-Caucasian individuals decreases. One of the main differences between phenotype frequencies of the Caucasian and African population is located in the so-called Duffy blood group system. This blood group system is similar to the ABO blood group system as it consists of six genotypes, four phenotypes $\mathcal{F} =$

(Fy(a-b-), Fy(a+b-), Fy(a-b+), Fy(a+b+)) and two antigens (Fy^a, and Fy^b). The phenotype frequencies for the Caucasian and African population are $f = (0, 0.18, 0.33, 0.49)$ and $f = (0.68, 0.06, 0.25, 0.01)$ respectively [94]. The probability that an African individual is Fy(a-b-) is 0.68 whereas the probability that a Caucasian individual has this combination is rare (< 0.0001). Hence, recruiting donors for this specific blood group randomly within the Caucasian population is virtually impossible. However, for a donor with this blood group combination the probability that a sibling has the same combination is 25% if the donor is Caucasian and 83% if the donor is African.

3.6   Conclusions

The generic mathematical approach described in this chapter allows computing a stationary genotype distribution for a given set of blood groups, which may even belong to multiple blood group systems. The input for the model consists of the phenotype distributions in a population only. This stationary genotype distribution allows answering a number of interesting questions using elementary statistics.

This chapter was tailored to quantify the effectiveness of targeted recruitment strategies aiming for relatives of donors with specific blood groups. It shows that the impact, in terms of the efficiency of targeting the next of kin of donors with known blood groups as potential new donors, can be substantial.

Recently, another application was found in computing the probability of a blood group mismatch between mother and fetus during pregnancy. This analysis also required an estimate of the stationary distribution of genotypes as a basis for further calculations. The approach outlined in this chapter therefore seems promising for answering various question related to genetic counseling.

3.7 Appendix I

Sets	
$\mathcal{A}$	Set of antigens
$\mathcal{L}$	Set of alleles
$\mathcal{H}$	Set of haplotypes (index h)
$\mathcal{G}$	Set of genotypes, with cardinality $|\mathcal{G}| = m$ (index γ)
$\mathcal{F}$	Set of phenotype, with cardinality $|\mathcal{F}| = n$ (index φ)
Matrices	
$S \in \{0,1\}^{m \times n}$	Matrix mapping blood group genotypes to blood group phenotypes
$P \in \mathbb{R}^{m \times m \times m}$	Heredity matrix, with $P\left(\gamma_k \mid \gamma_i, \gamma_j\right)$ the probability that two parents with genotypes γ_i and γ_j conceive a child with genotype γ_k
$A \in \mathbb{R}^{(m+n) \times m}$	Matrix used for the linearization of a system of quadratic equations, $(A^\top = [\boldsymbol{x}^\top P - I \quad S^\top])$
$I \in \mathbb{R}^{m \times m}$	Identity matrix
$Q \in \mathbb{R}^{(m+n) \times (m+n)}$	Orthogonal matrix used for QR factorization, $Q = [Q_1 \quad Q_2]$, $Q_1 \in \mathbb{R}^{(m+n) \times m}$, $Q_2 \in \mathbb{R}^{(m+n) \times n}$
$R \in \mathbb{R}^{(m+n) \times m}$	Upper triangular matrix used for the QR decomposition, $R^\top = [R_1 \quad 0]$, $R_1 \in \mathbb{R}^{m \times m}$, $0 \in \mathbb{R}^{n \times m}$
Vectors	
$\boldsymbol{x} \in \mathbb{R}^{m \times 1}$	Column vector representing a stationary genotype distribution
$\boldsymbol{f} \in \mathbb{R}^{n \times 1}$	Column vector representing the distribution of phenotypes in the general population
$\boldsymbol{v}_h \in \mathbb{R}^{m \times 1}$	Column vector containing the probabilities that a parent transmits haplotype $h \in \mathcal{H}$ to the child
$\boldsymbol{b} \in \mathbb{R}^{(m+n) \times 1}$	Column vector used for the linearizion of a system of quadratic equations ($\boldsymbol{b}^\top = [\boldsymbol{0} \quad \boldsymbol{f}]$)
$\boldsymbol{r} \in \mathbb{R}^{(m+n) \times 1}$	Column vector for the residuals $\boldsymbol{r} = A\boldsymbol{x} - \boldsymbol{b}$

The symbols $\mathcal{H}'$, $\mathcal{G}'$, B, $\boldsymbol{e}_m$, π, and a are omitted here, as they will be defined and only used in Section 3.8 (Appendix II).

3.8   Appendix II

In this appendix we fix ourselves to one and the same male genotype distribution $\boldsymbol{x}$ (as to be conceived for one iteration step as $\boldsymbol{x} = \boldsymbol{x}^{(n-1)}$). We only include $\boldsymbol{x}$ in the representation of $\mathcal{H}'(\boldsymbol{x})$ and $\mathcal{G}'(\boldsymbol{x})$ as these can be explicitly different for different $\boldsymbol{x}$.

Now, for given $\boldsymbol{x}$, let $A \in \mathbb{R}^{(m+n)\times m}$ be defined as follows:

$$A = \left[\begin{array}{c} \boldsymbol{x}^\top P - I \\ S^\top \end{array} \right], \tag{3.32}$$

where $\boldsymbol{x} \in \mathbb{R}^{m\times 1}$ is the genotype distribution, $P \in \mathbb{R}^{m\times m\times m}$ is the heredity matrix, $I \in \mathbb{R}^{m\times m}$ is the identity matrix, and $S \in \{0,1\}^{m\times n}$ is the matrix mapping blood group genotypes to blood group phenotypes. It needs to be shown that A has full column rank. Let $B \in \mathbb{R}^{m\times m}$ be the matrix $(B^\top = \boldsymbol{x}^\top P)$, with

$$B_{jk} = \sum_{\gamma_i \in \mathcal{G}} P\left(\gamma_k \mid \gamma_i, \gamma_j\right) x_{\gamma_i}, \tag{3.33}$$

where B_{jk} can be interpreted as the probability that a mother with genotype γ_j conceives a child with genotype γ_k, given that the genotype distribution of the father equals $\boldsymbol{x}$. Clearly, B is a stochastic matrix.

We can now regard B as a transition matrix of a Markov chain with state space $\mathcal{G}$. Hence, purely abstractly, the outcomes of the Markov chain can be seen as keeping track of the genotype distribution of mothers, which are based on the heredity matrix P and a *fixed* genotype distribution $\boldsymbol{x}$. (For the simple RhD example, Table 3.5 shows Markov chains for different $\boldsymbol{x}$.)
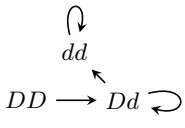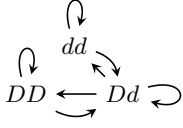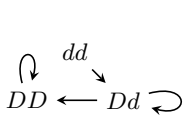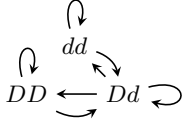
Let $\mathcal{H}'(\boldsymbol{x}) \subseteq \mathcal{H}$ be defined as the set of haplotypes that a father may transmit to his child, or differently stated: for every haplotype $h \in \mathcal{H}'(\boldsymbol{x})$ there exists a genotype $\gamma_i \in \mathcal{G}$, $\gamma_i = \{h, \cdot\}$ and/or $\gamma_i = \{\cdot, h\}$, such that $x_{\gamma_i} > 0$. This implies that the following two conditions should be satisfied:

$$\begin{cases} \sum\limits_{\gamma_i \in \mathcal{G}: h \in \gamma_i} x_{\gamma_i} > 0 & \text{if } h \in \mathcal{H}'(\boldsymbol{x}), \\ \sum\limits_{\gamma_i \in \mathcal{G}: h \in \gamma_i} x_{\gamma_i} = 0 & \text{if } h \notin \mathcal{H}'(\boldsymbol{x}). \end{cases} \tag{3.34}$$

As a consequence, roughly speaking, the haplotypes that are not present in the male population will eventually disappear. More precisely, the Markov chain, depending on the given genotype distribution $\boldsymbol{x}$, will always have a single closed class given by

$$\mathcal{G}'(\boldsymbol{x}) = \left\{ \gamma_k \in \mathcal{G} \mid \gamma_k = \{h_{k_1}, h_{k_2}\}, \quad h_{k_1}, h_{k_2} \in \mathcal{H}'(\boldsymbol{x}) \right\}. \tag{3.35}$$

Table 3.5: Markov chains for the RhD example.

$\boldsymbol{x}$	dd Dd DD $[1 \quad 0 \quad 0]$	dd Dd DD $[0 \quad 1 \quad 0]$	dd Dd DD $[0 \quad 0 \quad 1]$	dd Dd DD $[\frac{1}{2} \quad 0 \quad \frac{1}{2}]$
B	$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$
	dd $DD \longrightarrow Dd \circlearrowright$	dd $DD \rightleftarrows Dd \circlearrowright$	dd $DD \longleftarrow Dd \circlearrowright$	dd $DD \rightleftarrows Dd \circlearrowright$
$\mathcal{H}'(\boldsymbol{x})$	$\{d\}$	$\{d, D\}$	$\{D\}$	$\{d, D\}$
$\mathcal{G}'(\boldsymbol{x})$	$\{dd\}$	$\{dd, Dd, DD\}$	$\{DD\}$	$\{dd, Dd, DD\}$

Note that for $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, two *fixed* genotype distributions, and $a \in (0, 1)$ the following relation holds $\mathcal{H}'(a\boldsymbol{x}_1 + (1-a)\boldsymbol{x}_2) = \mathcal{H}'(\boldsymbol{x}_1) \cup \mathcal{H}'(\boldsymbol{x}_2)$, but $\mathcal{G}'(a\boldsymbol{x}_1 + (1-a)\boldsymbol{x}_2) = \mathcal{G}'(\boldsymbol{x}_1) \cup \mathcal{G}'(\boldsymbol{x}_2)$ is generally not true (see also Table 3.5).

Since the Markov chain has a single closed class, there exists a unique stationary distribution π satisfying

$$\begin{cases} B^\top \pi = \pi \\ \boldsymbol{e}_m^\top \pi = 1 \end{cases} \quad \Rightarrow \quad \begin{bmatrix} B^\top - I \\ \boldsymbol{e}_m^\top \end{bmatrix} \pi = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix},$$

where $\boldsymbol{e}_m$ is the all ones vector of length m. What remains, is to use this statement to proof the full column rank property of A.

The only candidate vector $\boldsymbol{y}$ (up to a constant), such that $A\boldsymbol{y} = \boldsymbol{0}$, is $\boldsymbol{y} = \boldsymbol{\pi}$, the stationary distribution of the matrix B. Now note that column sum of each column of $S \in \{0, 1\}^{m \times n}$ is at least one (i.e., $S^\top \boldsymbol{e}_n \geq \boldsymbol{e}_m$) and therefore $S^\top \boldsymbol{y} = \boldsymbol{0}$ if $\boldsymbol{y} = \boldsymbol{0}$. Hence, the only vector satisfying $A\boldsymbol{y} = \boldsymbol{0}$ is $\boldsymbol{y} = \boldsymbol{0}$, which implies that the matrix A has full column rank.

Part III

Inventory allocation

Blood groups: a binary representation

This chapter is a preliminary chapter for Chapters 5, 6, and 7. It introduces a binary representation for general blood groups (i.e., beyond the ABO, Rhesus-D blood groups), which provides an efficient way to determine the compatibility between blood groups of donors and transfusion recipients. In addition, it presents how the blood group distribution can be computed for an arbitrary set of antigens.

4.1 Introduction

Currently, the International Society of Blood Transfusion (ISBT) recognizes 38 different blood group systems, containing more than 300 different blood group antigens [108]. These antigens are surface markers on the outside of the RBC membrane and determine the blood group of an individual. As each antigen is either present or absent, the number of different blood groups is theoretically at least

$$\underbrace{2 \times 2 \times 2 \times ... \times 2}_{300 \text{ times}} = 2^{300}. \tag{4.1}$$

To illustrate the magnitude of this number, we might compare it to the number of particles on earth and conclude that the number of possible blood groups exceeds the number of particles!

In practice, however, the importance for matching on a specific antigen depends on the immunogenicity of this antigen, which is defined as the ability to provoke an immune response in the human body. More specifically, exposure to RBCs carrying antigens not present on the RBCs of a recipient itself might elicit an immune response in some individuals. The recipient develops antibodies against the foreign antigens and becomes immunized (alloimmunization). These antibodies could destroy the donor's RBCs during a subsequent blood transfusion, if the donor's RBCs carry the cognate antigen. In a similar way, antibodies might also cause problems during pregnancy. Most antibodies are small enough to pass the placenta

Table 4.1: Blood group systems considered in this thesis and the corresponding antigens that belong to these blood groups systems. Antigens k, M, and N are excluded, as we regard them of minor clinical importance (high/low-incident antigens or antigens corresponding to cold antibodies).

Blood group system	ABO		Rhesus					Kell		Duffy		Kidd		MNS			
Antigens	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	A	B	D	C	c	E	e	K	k	Fy^a	Fy^b	Jk^a	Jk^b	M	N	S	s
Antigens excluded									*					*	*		

and enter the bloodstream of the fetus, where they can destroy the fetus's RBCs. Alloimmunization due to a blood transfusion can be prevented by selecting a donor, whose RBCs lack the antigens that are not present on the RBCs of the recipient. Though, alloimmunization as a result of pregnancy is hard to prevent. Only ABO incompatibility between mother and child and anti-D immunoglobulin are known to reduce the probability of becoming immunized. However, no immunoglobulins have been developed against non-D antigens [135].

The mathematical models presented in this thesis could theoretically be applied to any number of antigens. Nevertheless, the numerical experiments will be limited to the six major blood group systems containing the fourteen clinically most relevant antigens, as presented in Table 4.1. This set of fourteen antigens coincides with the antigens considered in the current blood transfusion practice, both in the Netherlands and worldwide.

The remainder of this chapter is organized as follows. In Section 4.2, we first introduce a binary representation for the well-known ABO blood groups, followed by a binary representation for general blood groups in Section 4.3. This binary representation provides a clear and unique way to represent blood groups, irrespective of the number of antigens considered. Finally, in Section 4.4, we compute the blood group distribution for an arbitrary set of antigens.

4.2   ABO blood groups

The ABO blood groups are determined by just two antigens, A and B, which are either present or absent on the surface of the individual's RBCs. With these two antigens four different blood groups can be formed: O, A, B, and AB. Each blood group can also be represented by a set, whose elements are the antigens that are present on the RBCs. The empty set { } refers to blood group O (lacking both antigens A and B), {A} refers to blood group A (lacking antigens B), {B} refers to blood group B (lacking antigens A), and {A, B} refers to blood group AB (lacking no antigens), see Table 4.2. The set of ABO blood groups can then be presented

Table 4.2: Two mathematical representations of the ABO blood groups: a set representation ($\mathcal{I}$) and a binary vector representation (i).

	O	A	B	AB
$\mathcal{I}$	{ }	{A}	{B}	{A, B}
$i^{\top}$	[0 0]	[1 0]	[0 1]	[1 1]

as

$$\mathcal{B} = \{\{\ \},\{A\},\{B\},\{A,B\}\}, \tag{4.2}$$

which is equal to the power set of the set of antigens, i.e., $\mathcal{B}$ consists of all possible subsets of the set $\{A, B\}$.

With respect to blood group matching, however, it is more convenient to describe blood groups by binary vectors instead of sets of antigens, as this simplifies the comparison of blood groups. To illustrate this binary vector representation, introduce for each blood group $\mathcal{I} \in \mathcal{B}$ a binary vector $i = [i_1\ i_2]^{\top}$ with

$$i_1 = \begin{cases} 1 & \text{if } A \in \mathcal{I}, \\ 0 & \text{if } A \notin \mathcal{I}, \end{cases} \quad \text{and} \quad i_2 = \begin{cases} 1 & \text{if } B \in \mathcal{I}, \\ 0 & \text{if } B \notin \mathcal{I}. \end{cases} \tag{4.3}$$

Note that each binary vector corresponds to a unique decimal number (i.e., the binary vector representations of blood groups O, A, B, and AB refer to the binary numbers 00, 01, 10, and 11 respectively, which are equal to the decimal numbers 0, 1, 2, and 3). The ABO blood groups can thus be ordered according to their binary vector representation.

4.2.1 ABO compatibility

The general principle of blood group matching is based on the compatibility between the blood groups of donors and transfusion recipients. Let $i = [i_1\ i_2]^{\top}$ be the blood group of a donor and let $j = [j_1\ j_2]^{\top}$ be the blood group of a transfusion recipient. Their blood groups are called compatible if all antigens that are present on the RBCs of the donor are also present on the RBCs of the recipient (donor to recipient), or similarly, all antigens that are absent on the RBCs of the recipient are also absent on the RBCs of the donor (recipient to donor), i.e.,

$$j_k = \begin{cases} 1 & \text{if } i_k = 1, \\ 0 \text{ or } 1 & \text{if } i_k = 0, \end{cases} \quad \text{and} \quad i_k = \begin{cases} 0 \text{ or } 1 & \text{if } j_k = 1, \\ 0 & \text{if } j_k = 0, \end{cases} \tag{4.4}$$

with $k \in \{1, 2\}$. Hence, the compatibility between blood groups of donors and transfusion recipients can be determined by an element-wise comparison of two binary vectors, as elaborated in Table 4.3a. The traditionally used compatibility matrix, presented in Table 4.3b, is thus no longer required.

Table 4.3: Two ways to determine the compatibility between ABO blood groups: by an element-wise comparison of binary vectors ($i \leq j$) and by using a compatibility matrix (C_{ij}).

(a) Element-wise comparison of binary vectors.

	$j^\top$ (donor to recipient)	$i^\top$ (recipient to donor)
[0 0]	[0 0], [1 0], [0 1], [1 1]	[0 0]
[1 0]	[1 0], [1 1]	[0 0], [1 0]
[0 1]	[0 1], [1 1]	[0 0], [0 1]
[1 1]	[1 1]	[0 0], [1 0], [0 1], [1 1]

(b) Compatibility matrix.

$i^\top / j^\top$	[0 0]	[1 0]	[0 1]	[1 1]
[0 0]	1	1	1	1
[1 0]	0	1	0	1
[0 1]	0	0	1	1
[1 1]	0	0	0	1

Remark 4.1 (Blood groups). In the remainder of this thesis, the vectors i and j are reserved for the blood groups of the donor and transfusion recipient, respectively, i.e.,

i: blood group donor,
j: blood group recipient.

4.3   General blood groups

To represent general blood groups, let $\mathcal{A} = \{a_1, a_2, ..., a_n\}$ be a finite set of antigens and let $\mathcal{B} = \{\mathcal{I}_1, ..., \mathcal{I}_m\}$ be a finite set of blood groups, where $n = |\mathcal{A}|$ represents the number of antigens considered and $m = |\mathcal{B}|$ represents the number of blood groups considered. A blood group $\mathcal{I} \in \mathcal{B}$ is defined as a unique combination of antigens $a_k \in \mathcal{A}$ that are either present ($a_k \in \mathcal{I}$) or absent ($a_k \notin \mathcal{I}$) on the surface of an RBC. An individual with blood group $\mathcal{I} \in \mathcal{B}$ is called a_k-positive if antigen $a_k \in \mathcal{A}$ is present on the individual's RBCs and a_k-negative if this antigen is absent. Due to the dichotomous structure of an antigen (i.e., it is either present or absent) the set of blood groups $\mathcal{B}$ that can be composed from antigens $a_k \in \mathcal{A}$ is equal to the power set of $\mathcal{A}$:

$$\mathcal{B} = \{\{\ \}, \{a_1\}, \{a_2\}, \{a_1, a_2\}, ..., \{a_1, .., a_n\}\}. \tag{4.5}$$

This implies that the cardinality of the set $\mathcal{B}$ grows exponentially with the number of antigens considered (i.e., $m = 2^n$), which may complicate computations when n is large.

4.3.1   Binary representation

From a mathematical and computational point of view, it is more convenient to describe a blood group $\mathcal{I} \in \mathcal{B}$ by a binary vector instead of a set of antigens. This

largely simplifies the mathematical representation, as already illustrated in Section 4.2 for the ABO blood groups. Therefore, for each blood group $\mathcal{I} \in \mathcal{B}$ a binary vector $i = [i_1 \ i_2 \ \cdots \ i_n]^\top$ is introduced with

$$i_k = \begin{cases} 1, & \text{if } a_k \in \mathcal{I} \ (\text{individual is } a_k\text{-positive}), \\ 0, & \text{if } a_k \notin \mathcal{I} \ (\text{individual is } a_k\text{-negative}), \end{cases} \tag{4.6}$$

for all $k \in \{1, ..., n\}$. This binary representation of a blood group is particularly useful to investigate whether two blood groups are compatible, but also provides an intelligent way to order them, as specified in Section 4.3.2.

Remark 4.2 (Set of blood groups). As each set of antigens $\mathcal{I} \in \mathcal{B}$ corresponds to a unique binary vector $i \in \{0,1\}^n$, in the remainder of this thesis the set $\mathcal{B}$ is used to refer to a set of blood groups, whose elements can be either sets of antigens ($\mathcal{I} \in \mathcal{B}$) or binary vectors ($i \in \mathcal{B}$).

4.3.2   Ordered set of blood groups

Colexicographical order

The set of binary vectors of length n can be ordered in multiple ways, of which the lexicographical order would probably be the most obvious one. Nevertheless, we prefer to order them differently, into a colexicographical order. The colexicographical order (expressed as '$\preceq$') formally states that a vector $i \in \mathcal{B}$ precedes a vector $j \in \mathcal{B}$ if for some $k \in \{1, ..., n\}$ it holds that $i_k < j_k$ and $i_\ell = j_\ell$ for all $\ell > k$. The colexicographically ordered set of binary vectors of length n is therefore given by

$$(\mathcal{B}, \preceq) = \left\{ \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, ..., \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\} \quad \begin{matrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{matrix} \quad . \tag{4.7}$$
$$\quad\quad\quad\quad\ 0 \quad\ 1 \quad\ 2 \quad\ 3 \quad\ 2^n - 1$$

The same order can be obtained by transforming each binary vector $i = [i_1 \ i_2 \ \cdots \ i_n]^\top$ into a decimal number and ordering these decimal numbers from 0 to $2^n - 1$.

Product order

In contrast to the colexicographical order, which belongs to the class of total orders, the so-called product order belongs to the class of partial orders, implying that there might not exist a precedence relation between some pairs of vectors $i, j \in \mathcal{B}$. The product order (expressed as '$\leq$') formally states that
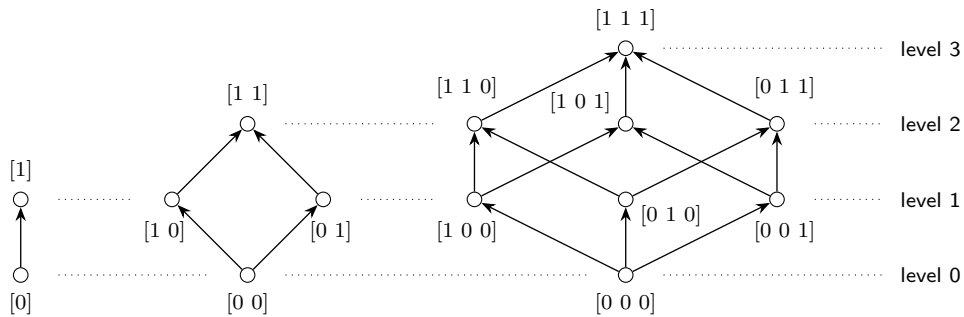
Figure 4.1: Hasse diagram of the partially ordered set of blood groups $(\mathcal{B}, \leq)$ for $n = 1, 2, 3$.

- i precedes (or coincides with) j, if $i_k \leq j_k$ for all $k \in \{1, ..., n\}$,

- j precedes (or coincides with) i, if $j_k \leq i_k$ for all $k \in \{1, ..., n\}$,

- i and j are incomparable otherwise,

for each pair of vectors $i, j \in \mathcal{B}$. The partially ordered set $(\mathcal{B}, \leq)$ can be visualized by a Hasse diagram; an upward oriented directed graph with vertex set $\mathcal{B}$ and edge set $\{[i, j] \in \mathcal{B} \times \mathcal{B} \mid i \leq j, \ \|i\|_1 + 1 = \|j\|_1\}$, as illustrated in Figure 4.1 for $n = 1, 2, 3$. In other words, each vector $i \in \mathcal{B}$ is assigned to a level based on the number of ones appearing in this vector, and comparable vectors are connected by an edge if they are assigned to consecutive levels.

4.3.3   Blood group compatibility

The partially ordered set $(\mathcal{B}, \leq)$ and the corresponding Hasse diagram appear to be perfectly suited to describe the compatibility between the blood groups of donors and transfusion recipients. A match between a donor with blood group $i \in \mathcal{B}$ and a transfusion recipient with blood group $j \in \mathcal{B}$ is called

- identical, if all antigens present on the RBCs of the donor are also present on the RBCs of the recipient and all antigens absent on the RBCs of the donor are also absent on the RBCs of the recipient (i.e., $i_k = j_k$ for all $k \in \{1, ..., n\}$),

- compatible, if all antigens present on the RBCs of the donor are also present on the RBCs of the recipient, or similarly, if all antigens absent on the RBCs of the recipient are also absent on the RBCs of the donor (i.e., $i_k \leq j_k$ for all $k \in \{1, ..., n\}$),

- incompatible, if there exists at least one antigen present on the RBCs of the donor, but absent on the RBCs of the transfusion recipient (i.e., $i_k > j_k$ for some $k \in \{1, ..., n\}$).

Remark 4.3 (Compatibility matrix). The compatibility between the blood groups of a donor with blood group $i \in \mathcal{B}$ and transfusion recipient with blood group $j \in \mathcal{B}$ can be determined by an element-wise comparison of these binary vectors. The matrix $C \in \{0, 1\}^{m \times m}$, with

$$C_{ij} = \begin{cases} 1 & \text{if } i \leq j, \\ 0 & \text{otherwise,} \end{cases} \tag{4.8}$$

for all $i, j \in \mathcal{B}$, which was previously used to model the compatibility between blood groups [29, 35, 56, 60, 132], is no longer required. This substantially reduces the computational complexity when general blood groups are considered, as the number of elements in this matrix grows exponentially with the number of antigens considered (i.e., $m^2 = 4^n$).

4.4 Blood group distribution

As stated in Section 4.3, a blood group $\mathcal{I} \in \mathcal{B}$ is defined as a unique combination of antigens $a \in \mathcal{A}$ that are either present or absent on the surface of the RBCs. Let $f_a(\mathcal{A})$ indicate the probability that a particular antigen is present and $1 - f_a(\mathcal{A})$ indicate the probability that a particular antigen is absent, then the relative frequency of a blood group $\mathcal{I} \in \mathcal{B}$ would simply be a product of these probabilities, i.e.,

$$f_{\mathcal{I}}(\mathcal{A}) \neq \left(\prod_{a \in \mathcal{I}} f_a(\mathcal{A}) \right) \cdot \left(\prod_{a \notin \mathcal{I}} (1 - f_a(\mathcal{A})) \right) \tag{4.9}$$

if these probabilities are mutually independent. This relation of mutual independence does, however, not hold for antigens belonging to the same blood group system. The distribution of antigens within a blood group system has, therefore, first to be determined for each system separately. These antigen distributions can be found in [94] (for convenience also included in Section 4.5, Appendix I), but still have to be aggregated based upon the specific set of antigens considered.

Let $\mathcal{S}^k$ be a set containing the antigens that belong to blood group system k. Define $\mathcal{A}^k = \mathcal{A} \cap \mathcal{S}^k$ and $\mathcal{I}^k = \mathcal{I} \cap \mathcal{S}^k$. The sets $\mathcal{A}$ and $\mathcal{I}$ are both partitioned into mutually disjoint sets, such that $\mathcal{A} = \cup_k \mathcal{A}^k$ and $\mathcal{I} = \cup_k \mathcal{I}^k$. The relative frequency of $\mathcal{I}^k$ can be computed by the following formula:

$$f_{\mathcal{I}^k}(\mathcal{A}^k) = \sum_{\mathcal{L} \in \mathcal{U}^k} f_{\mathcal{L}}(\mathcal{S}^k) \tag{4.10}$$

where $\mathcal{U}^k = \left\{ \mathcal{I}^k \cup \mathcal{T} \mid \mathcal{T} \in \mathcal{P}\left\{\mathcal{S}^k \backslash \mathcal{A}^k\right\}\right\}$. Finally, as antigens between blood group systems are mutually independent, the relative frequencies of antigen combinations within blood group system are multiplied to obtain the relative frequency for blood group $\mathcal{I} \in \mathcal{B}$, i.e.,

$$f_{\mathcal{I}}(\mathcal{A}) = \prod_k f_{\mathcal{I}^k}(\mathcal{A}^k), \tag{4.11}$$

as elaborated in Example 4.1.

These formulas will thoroughly be used in Chapter 7 as to compute antigen distributions for more than a thousand different matching strategies.

Example 4.1 (Relative frequency of a blood group). Consider the following set of antigens $\mathcal{A}$ and let a specific blood group $\mathcal{I} \in \mathcal{P}(\mathcal{A})$ be given by:

$$\mathcal{A} = \{ \underbrace{\{\mathsf{A}, \mathsf{B}\}}_{\mathcal{A}_1} \mid \underbrace{\{\mathsf{D}, \mathsf{C}, \mathsf{E}\}}_{\mathcal{A}_2} \mid \underbrace{\{\mathsf{K}\}}_{\mathcal{A}_3} \mid \underbrace{\{\mathsf{Fy}^{\mathsf{a}}\}}_{\mathcal{A}_4} \mid \underbrace{\{\ \}}_{\mathcal{A}_5} \mid \underbrace{\{\mathsf{S}, \mathsf{s}\}}_{\mathcal{A}_6} \},$$

$$\mathcal{I} = \{ \underbrace{\{\ \}}_{\mathcal{I}_1} \mid \underbrace{\{\mathsf{D}, \mathsf{E}\}}_{\mathcal{I}_2} \mid \underbrace{\{\ \}}_{\mathcal{I}_3} \mid \underbrace{\{\mathsf{Fy}^{\mathsf{a}}\}}_{\mathcal{I}_4} \mid \underbrace{\{\ \}}_{\mathcal{I}_5} \mid \underbrace{\{\mathsf{s}\}}_{\mathcal{I}_6} \}$$

respectively, which are both partitioned into six sets, one for each blood group system. The relative frequency of blood group $\mathcal{I}$ is equal to

$$
\begin{aligned}
f_{\mathcal{I}}(\mathcal{A}) &= f_{\mathcal{I}^1}(\mathcal{A}^1) \times f_{\mathcal{I}^2}(\mathcal{A}^2) \times f_{\mathcal{I}^3}(\mathcal{A}^3) \times f_{\mathcal{I}^4}(\mathcal{A}^4) \times f_{\mathcal{I}^5}(\mathcal{A}^5) \times f_{\mathcal{I}^6}(\mathcal{A}^6) \\
&= \quad 0.44 \quad \times \quad 0.14 \quad \times \quad 0.91 \quad \times \quad 0.66 \quad \times \quad 1 \quad \times \quad 0.45 \\
&= \quad 0.017,
\end{aligned}
$$

where $f_{\mathcal{I}^k}(\mathcal{A}^k)$, $k \in \{1, ..., 6\}$, is computed by equation (4.10).

Remark 4.4 (Blood group distribution). In this thesis, we assume that the distribution of blood groups in the donor and transfusion recipient population are both equal to the distribution of blood groups in a Caucasian population (a population consisting of individuals with European ancestors) [94]. However, all mathematical models presented can also be applied for other blood group distributions. Moreover, the distribution of blood groups in the donor and transfusion recipient population might actually differ.

4.5   Appendix I

This appendix contains antigen distributions for the six blood group systems considered in this thesis: ABO, Rhesus, Kell, Duffy, Kidd, and MNS. Note that for the Rhesus and MNS blood group systems, not all antigen combinations are included, since these do not exsist and are, therefore, also not included in [94]. Antigen combinations classified as 'rare' are assumed to have a relative frequency of zero.

Table 4.4: Relative frequency of antigen combinations in a Caucasian population for the ABO, Rhesus, Kell, Duffy, Kidd, and MNS blood group systems. This table is based on tables presented in [94], but extended with the columns 'set representation' and 'binary representation'.

Blood group system	Specific name	Set representation	Binary representation	Relative frequency
ABO	O	{ }	[0 0]	0.44
	A	{A}	[1 0]	0.43
	B	{B}	[0 1]	0.09
	AB	{A, B}	[1 1]	0.04
Rhesus	rr	{c, e}	[0 0 1 0 1]	0.15
	r"r"	{c, E}	[0 0 1 1 0]	rare
	r"r	{c, E, e}	[0 0 1 1 1]	0.01
	r'r'	{C, e}	[0 1 0 0 1]	rare
	$r^y r^y$	{C, E}	[0 1 0 1 0]	rare
	$r'r^y$	{C, E, e}	[0 1 0 1 1]	rare
	r'r	{C, c, e}	[0 1 1 0 1]	0.01
	$r"r^y$	{C, c, E}	[0 1 1 1 0]	rare
	r'r"	{C, c, E, e}	[0 1 1 1 1]	rare
	$R_0 r$	{D, c, e}	[1 0 1 0 1]	0.02
	$R_2 R_2$	{D, c, E}	[1 0 1 1 0]	0.02
	$R_2 r$	{D, c, E, e}	[1 0 1 1 1]	0.12
	$R_1 R_1$	{D, C, e}	[1 1 0 0 1]	0.19
	$R_z R_z$	{D, C, E}	[1 1 0 1 0]	rare
	$R_1 R_z$	{D, C, E, e}	[1 1 0 1 1]	0.00
	$R_1 r$	{D, C, c, e}	[1 1 1 0 1]	0.35
	$R_2 R_z$	{D, C, c, E}	[1 1 1 1 0]	0.00
	$R_1 R_2$	{D, C, c, E, e}	[1 1 1 1 1]	0.13
Kell		{ }	[0 0]	rare
		{K}	[1 0]	0.00
		{k}	[0 1]	0.91
		{K, k}	[1 1]	0.09

Blood group system	Specific name	Set representation	Binary representation	Relative frequency
Duffy		{ }	[0 0]	rare
		$\{Fy^a\}$	[1 0]	0.17
		$\{Fy^b\}$	[0 1]	0.34
		$\{Fy^a, Fy^b\}$	[1 1]	0.49
Kidd		{ }	[0 0]	rare
		$\{Jk^a\}$	[1 0]	0.26
		$\{Jk^b\}$	[0 1]	0.23
		$\{Jk^a, Jk^b\}$	[1 1]	0.50
MNS		{M}	[1 0 0 0]	rare
		{N}	[0 1 0 0]	rare
		{M, N}	[1 1 0 0]	rare
		{M, S}	[1 0 1 0]	0.06
		{N, S}	[0 1 1 0]	0.01
		{M, N, S}	[1 1 1 0]	0.04
		{M, s}	[1 0 0 1]	0.08
		{N, s}	[0 1 0 1]	0.15
		{M, N, s}	[1 1 0 1]	0.22
		{M, S, s}	[1 0 1 1]	0.14
		{N, S, s}	[0 1 1 1]	0.06
		{M, N, S, s}	[1 1 1 1]	0.24

A microscopic mathematical description for optimal blood issuing

5.1 Introduction

5.1.1 Motivation

Blood service organizations collect blood from donors and, after rigorous testing and processing, fractionate it into red blood cells (RBCs), platelets and plasma. RBCs and platelets have limited shelf lives and have to be discarded once passing their shelf lives. A blood bank is responsible for an adequate supply of blood (blood components) to hospitals. A shortage in supply may put the lives of patients at risk. Hence, blood banks usually keep a sizable safety stock to avoid such an event. However, as RBC units are perishable, too much stock on hand increases outdating imposing additional costs to the system. Moreover, as these blood components are drawn from voluntary non-remunerated donors, whose main motivation is to help patients, outdating should be prevented.

Blood banks issue items according to hospital orders. Hospitals do not specify the age or the distribution of age of items in their orders, however, receiving fresher RBC units reduces outdating at hospitals. In addition, recent studies show that transfusing fresh blood may lead to an improvement of patient outcomes and prospective studies are currently being performed to check whether fresh RBCs are clinically preferred for certain patient groups [39].

5.1.2 Literature

Inventory management of perishable products is a well-studied topic within the field of Operations Research with over a hundred papers in the last decades, see for example [4, 50, 62, 79] for a review on this topic. Also the inventory management of blood products garners much attention itself and in the last years three review papers about this topic appeared [7, 83, 86].

Literature on inventory management of perishable products often assumes that replenishment is controllable, which implies that products can be produced, or in case of a blood bank, donors can be invited to donate. Moreover, to prevent outdating items are often issued according to a FIFO policy. In fact, a FIFO issuing policy does not only prevent outdating, but also shortages as the probability that the inventory runs empty decreases when outdating decreases.

An important difference between inventory management of RBC units and other perishable products is that the replenishment is less controllable, because the inputs to the inventory (i.e. donations) are stochastic. To consider the stochastic behavior in the blood inventory system, it can be regarded as a queuing system.

In most of these queuing models, it is assumed that fresh RBC units (donations) and requests arrive, one-by-one, according to independent Poisson processes with rates λ and μ, respectively. The maximum shelf life of an RBC unit is equal to a finite constant R. Requests are either satisfied according to a predetermined issuing policy (i.e., FIFO, LIFO, threshold) or lost when the inventory is empty. Under these assumptions, the inventory of RBC units can be regarded as an $M/M/1+D$ queue, where the $+D$ at the end of Kendall's notation (usually referring to the distribution of customer impatience) stands for the deterministic maximum shelf life of an RBC unit. Hence, RBC units that are not issued within R days expire and are removed from the inventory system.

As stated in Sarhangian et al. [101], the assumption of one-by-one arrivals of donations and requests might seem restrictive and unrealistic in practice. Blood collection sites transport only once or twice a day the donated whole blood units to a production center (where they are fractionated into RBC units and other blood products). Hospitals often request multiple RBC units at the same time to replenish their inventories. Though, Sarhangian et al. [101] also showed that an $M/M/1+D$ queuing model provides a very accurate approximation for systems with Poisson distributed batch donations and Poisson distributed batch requests. Even for non-Poisson distributed batch arrivals the qualitative observations remained valid.

The queuing model of interest was first introduced by Graves [51], who determined the stationary distribution of the age of the oldest RBC unit in inventory under a FIFO issuing policy. This stationary distribution was used to provide analytical expressions for performance measures of the inventory system, i.e., the probability of shortage, the probability of outdating, the average age of issued RBC units, and the average number of RBC units in inventory. The $M/M/1+D$ queuing model under a FIFO issuing policy was further studied by Kaspi and Perry [64], Parlar et al. [84], and Sarhangian et al. [101]. They analyzed the so-called virtual outdating process, which returns the remaining time until the next outdate occurs, and determined the distributions of the number of RBC units outdated in a busy period and the number of RBC units short in an idle period.

5.1.3 Contributions

The main contribution of this chapter to the existing literature is that it provides a general description and formulation for issuing RBC units upon requests. This formulation is independent of the number of antigens and blood groups considered and can thus be applied for general blood groups. In contrast to the papers presented, we take advantage of the fact that donations and requests arrive according to independent Poisson processes. We approximate the deterministic maximum shelf life of an RBC unit by an Erlang distribution. Hence, following Kendall's classification, this process can be referred to as an $M/M/1 + $ Erlang queue, which is equal to a continuous-time Markov chain when the age of the RBC units is part of the state description. We show that under a fixed issuing policy (i.e., FIFO, LIFO, threshold) the stationary distribution of the inventory can be computed by a so-called product form. Moreover, we extend this continuous-time Markov chain to a continuous-time Markov decision chain by including actions about which RBC unit to issue.

The remainder of this chapter is organized as follows. First, in Section 5.2, we discuss preliminary results on blood inventory allocation problem itself. Next, in Section 5.3, we show that the ageing process of a single RBC unit can be approximated by a phase-type distribution. Then, in Section 5.4 the inventory allocation problem, and its structure as a GSMP with actions, will be formulated as a Markov decision process.

5.2 Preliminaries

The main purpose of this chapter is to transform the blood inventory allocation problem of which RBC unit to issue for two reasons:

1) to reduce all possible transitions to a small number of neighbouring transitions

2) to come to an iterative and replicable computational scheme.

5.2.1 Inventory allocation problem

This subsection provides a global description for the blood inventory allocation problem of interest. It behaves like a continuous-time stochastic discrete event system (with decisions) and describes how the composition of the inventory evolves over time. For the time being, we assume a FIFO issuing policy and do not distinguish between different blood groups, implying that the decision about which RBC unit to issue as to satisfy a particular request is fixed. Later, in Section 5.4 we lift this assumption and elaborate on the decision dimension.

Assume that donations arrive according to a Poisson process with rate λ and each donation consists of a single RBC unit. Similarly, assume that requests occur according to a Poisson process with rate μ and each request consists of a single RBC unit. When items arrive into inventory, there residual shelf life is set equal to R. Items that remain in inventory for the entire R time units expire and are removed from inventory. If no RBC units are present in inventory, when a request arrives, a shortage occurs. Moreover, no backorders are permitted.

A state, defined as a particular composition of the inventory, can be seen as a snapshot of the inventory at a specific point in time. It indicates the number of RBC units in inventory, their blood groups, and their residual shelf lives. Depending on the state of the inventory three possible events can be distinguished:

- a fresh RBC unit is donated (with rate λ),

    - due to which the number of RBC units in inventory increases by one,

- an RBC unit is requested (with rate μ),

    - due to which the number of RBC units in inventory either decreases by one (request satisfied),

    - or remains unchanged (request is not satisfied incurring a shortage),

- all RBC units age (linearly with the time spend in inventory),

    - due to which the number of RBC units in inventory remains unchanged

    - until the residual shelf life of an RBC unit becomes zero, at which point this unit is removed from inventory, Hence, the number of RBC units decreases by one.

An action indicates which RBC unit should be selected from inventory to satisfy a request for an RBC unit with a particular blood group or whether it should be rejected (introducing a shortage). The system that will be studied in this Chapter is shown in Figure 5.1.

5.2.2   A Markovian overview

As illustrated by Figure 5.1, at each moment multiple times or clocks can be running simultaneously: one for the exponential time up to a new donation, one for the exponential time until a new request, and multiple deterministic ones that evolve entirely parallel (linear with time) to reflect ageing for each RBC unit in inventory up to its outdating. This naturally reflects the structure of a so-called generalized semi-Markov process (GSMP).

Standardly, the concept of a Markov process in discrete-time and continuous-time are well-known in literature. Loosely speaking, a stochastic process is said to

Figure 5.1: Continuous-time stochastic discrete event system for the inventory allocation problem. RBC units are donated according to a Poisson process with rate λ (circles). RBC units are requested according to a Poisson process with rate μ (squares). Upon arrival of a request either the oldest RBC unit is issued from inventory to satisfy this request (green square) or no RBC unit is selected, incurring a shortage (red square). Finally, the RBC units in inventory ages until they are either issued or expire
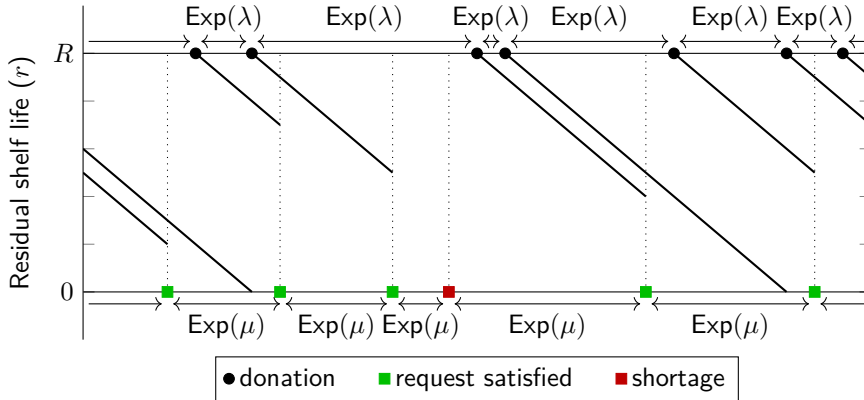
be Markovian, and is formally called a Markov process, if the probability distribution of future states depends only on the present state, independently of the states that have been visited in the past. This Markovian property appears to apply to many real-world systems that evolve dynamically in time, including inventory systems. In discrete-time and continuous-time Markov chains, there is only a single time component, up to a next state change. This can be

- a fixed time unit (for discrete-time Markov chains),

- an exponential time duration (for continuous-time Markov chains),

- or even a non-exponential time duration (for semi-Markov processes)

The distinctive character of a GSMP is that multiple non-exponential durations can begin or end at different times. As we will not be using these concepts explicitly in this chapter, we only give a brief formal description of a DTMC and a GSMP in Section 5.6, Appendix I. With reference to Remark 5.1, in Figure 5.2 we just present a schematic overview for conceptualization. In addition, as in line with the inventory allocation problem, in this figure we also incorporate the possibility of decisions upon a transition.

In the subsequent sections, the inventory allocation problem will be specified in a self-contained manner.

Remark 5.1. Figure 5.2 is intended to just present a schematic overview for structuring this chapter. By no means, does it aims to be fully defining. In fact, as

it turns out after thorough inspection of literature, no unique or fully distinctive consistent use of the words chain or process are generally found. Other than the "Markovian property" to relate to the history independence at a given state, the word "decision", in case actions can be involved, and the word "semi-", in case inter-jump times are not given by a fixed time unit (in discrete time) or one or multiple (general) exponential times (in continuous time).

5.3   Ageing of RBC units

The GSMP, as described in Section 5.2.1, can be transformed into a continuous-time Markov chain by approximating each non-exponential delay distribution by a continuous phase-type distribution [80, 91, 131]. The only non-exponential delay distribution in the GSMP concerns the maximum time that an RBC unit can be kept in inventory before it expires, which is given by a known parameter R.

First, in Section 5.3.1, a single RBC unit is considered. We argue that the maximum residual shelf life of an RBC unit can be approximated by a phase-type distribution, or more precisely, by an Erlang distribution. Next, in Section 5.3.2, the same approximation is expanded to multiple RBC units. The aggregated phase-type distributions can be regarded as infinite server tandem queues, used to describe how the composition of the inventory evolves over time.

Figure 5.2: Relation between a Markov chain (MC), Markov process (MP), semi-Markov process (SMP), generalized semi-Markov process (GSMP), and their decisional counter parts: discrete / continuous-time Markov decision process (MDP), semi-Markov decision process (SMDP), and generalized semi-Markov decision process (GSMDP). Referring to Remark 5.3 below, this figure is inspired by figures in references [91, 131], but it has not been presented in this specific way.
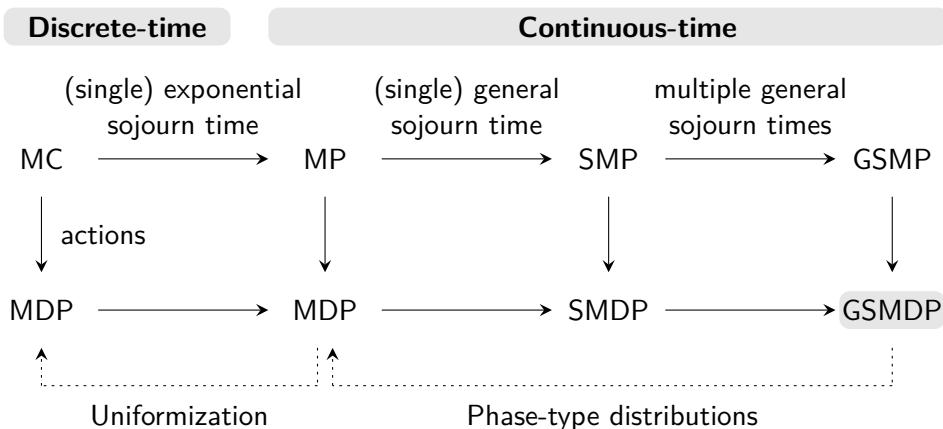
5.3.1 Single RBC unit

The ageing process (or residual shelf life) of a single RBC unit can formally be described as a deterministic continuous-time process $\rho = \{\rho(t) \mid t \geq 0\}$, with

$$\rho(t) = \begin{cases} R - t & \text{if } 0 \leq t < R, \\ 0 & \text{if } t \geq R, \end{cases} \tag{5.1}$$

where R is a known parameter, representing the maximum residual shelf life of an RBC unit. In other words, when a fresh RBC unit enters the inventory its residual shelf life is set equal to R. Its residual shelf life decreases linearly with the time spent in inventory. If the RBC unit is not issued within R days, it expires and is removed from inventory.

Deterministic approximation

In practice, it is realistic to discretize the residual shelf life of an RBC unit to a more natural counting parameter, i.e., the number of weeks or days left until expiration, as illustrated in Figures 5.3a and 5.3b respectively. The ageing process can then be regarded as a deterministic continuous-time (jump) process $\hat{\rho} = \{\hat{\rho}(t) \mid t \geq 0\}$, with

$$\hat{\rho}(t) = \begin{cases} R - \frac{(k-1)R}{N} & \text{if } t_{k-1} \leq t < t_k, \\ 0 & \text{if } t \geq t_N, \end{cases} \tag{5.2}$$

where $t_k = \frac{kR}{N}$, $k \in \{0, 1, ..., N\}$ are fixed time moments at which jumps occur and N is the number of segments in which the time-interval $[0, R]$ is divided. The time between two consecutive jumps $\tau_k = t_k - t_{k-1}$ is constant, implying that the length of each segment $[t_{k-1}, t_k)$ is deterministic. Moreover, since $\rho(t)$ is linear in t, the jump size $\hat{\rho}(t_k) - \hat{\rho}(t_{k-1})$ is also a constant. The continuous-time jump process $\hat{\rho}$ converges to ρ when the number of jumps N tends to infinity.

A disadvantage of this deterministic approximation is that it cannot be aligned with the exponential behavior of the other two events (i.e., donations and requests arrive according to independent Poisson processes). Therefore, instead of approximating the ageing process of an RBC unit by a deterministic time grid, it can also be approximated by an exponentially thinned time grid.

Stochastic approximation

Assume that the time between two consecutive jumps τ_k is exponentially distributed with mean $\nu_N^{-1} = R/N$. The ageing process can then be regarded as a continuous-time jump process $\tilde{\rho} = \{\tilde{\rho}(t) \mid t \geq 0\}$, with

$$\tilde{\rho}(t) = \begin{cases} R - \frac{(k-1)R}{N} & \text{if } T_{k-1} \leq t < T_k, \\ 0 & \text{if } t \geq T_N, \end{cases} \tag{5.3}$$

Figure 5.3: The ageing process of a single RBC unit can be approximated by a deterministic continuous-time (jump) process $\hat{\rho}$, consisting of N jumps, such that both the time between two consecutive jumps and the jump size are constant ($\tau_k = R/N$ and $\hat{\rho}(t_k) - \hat{\rho}(t_{k-1}) = -R/N$).
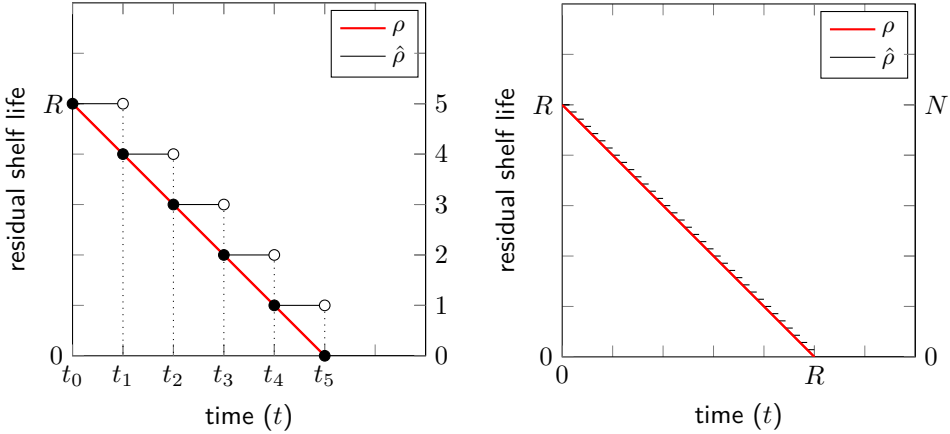
(a) Residual shelf life in weeks ($N = 5$) **(b)** Residual shelf life in days ($N = 35$)

where T_k, $k \in \{0, ..., N\}$, are stochastic time moments at which jumps occur, equal to the sum of k independent identically distributed exponential random variables, as illustrated in Figure 5.4. The maximum shelf life of an RBC unit, denoted by T_N, is thus no longer equal to the known parameter R, but approximated by an Erlang distribution with parameters N and ν_N. The expectation and variance of this distribution are given by

$$\mathbb{E}\left(T_N\right) = N\nu_N^{-1} = R \qquad \text{and} \qquad \text{Var}\left(T_N\right) = N\nu_N^{-2} = R^2/N, \tag{5.4}$$

respectively. As N and ν_N can be taken arbitrarily large – keeping $N\nu_N^{-1} = R$ – the variance of T_N converges to zero. This implies that the Erlang distribution converges to a deterministic distribution when the number of jumps N tends to infinity, as stated more formally in Proposition 5.1.

Proposition 5.1. Let $X = \mathbb{P}\left(\rho(t) = 0\right)$ be a deterministic random variable with cumulative distribution function

$$F_X(t) = \begin{cases} 1 & \text{if } t \geq R, \\ 0 & \text{if } 0 \leq t < R, \end{cases} \tag{5.5}$$

and let $Y = \mathbb{P}\left(\tilde{\rho}(t) = 0\right)$ be an Erlang random variable with cumulative distribution function

$$F_Y(t) = 1 - \sum_{k=0}^{N-1} \frac{e^{-\nu_N t}\left(\nu_N t\right)^k}{k!} \qquad t \geq 0, \tag{5.6}$$

Figure 5.4: The ageing process of a single RBC unit can be approximated by a continuous-time jump process $\tilde{\rho}$, consisting of N jumps, such that the time between two consecutive jumps is exponentially distributed ($\tau_k \sim \text{Exp}(\nu_N)$) and the jump size is constant ($\tilde{\rho}(T_k) - \tilde{\rho}(T_{k-1}) = -R/N$).

(a) Residual shelf life in weeks ($N = 5$)   **(b)** Residual shelf life in days ($N = 35$)
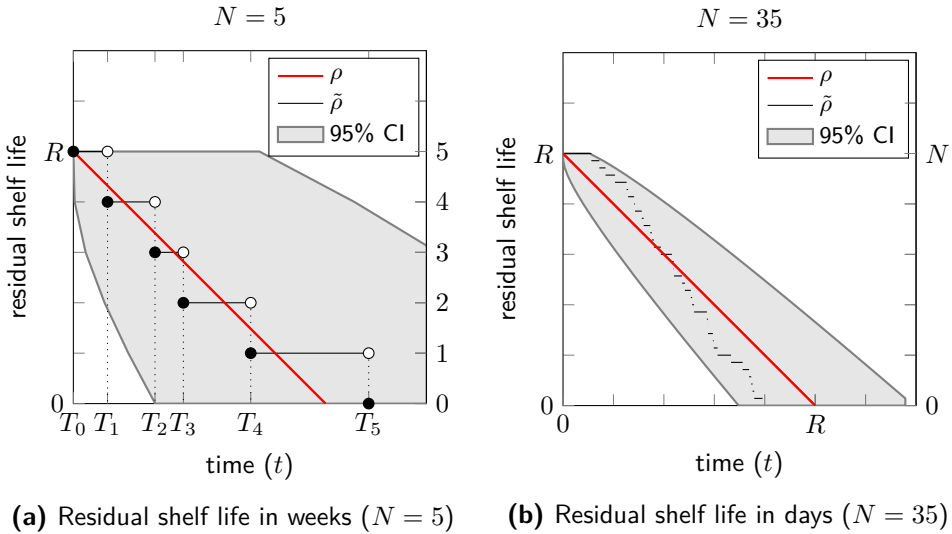
where $\nu_N = N/R$, then $\lim_{N\to\infty} F_Y(t) \xrightarrow{d} F_X(t)$ for all $t \geq 0$.

Proof. Observe that $F_Y(t) = \mathbb{P}(Z \geq N)$, where $Z \sim \text{Poisson}(\nu_N t)$. By applying Chebyshev's inequality, i.e., for any real number $k > 0$

$$\mathbb{P}(|Z - \nu_N t| \geq k) \leq \frac{\nu_N t}{k^2}, \tag{5.7}$$

it can be shown that

$$\begin{cases} F_Y(t) \geq 1 - N^{-1} \frac{Rt}{(R-t)^2} & \text{if } t > R, \\ F_Y(t) \leq N^{-1} \frac{Rt}{(R-t)^2} & \text{if } 0 \leq t < R. \end{cases} \tag{5.8}$$

These expressions converge to one and zero, respectively, when N tends to infinity. Hence, $\lim_{N\to\infty} F_Y(t) \xrightarrow{d} F_X(t)$ for all $t \geq 0$. $\qquad\square$

Note that the outcomes of the processes $\hat{\rho}$ and $\tilde{\rho}$ can be mapped to an integer $r \in \{N, N-1, ..., 1, 0\}$, by multiplying these outcomes by N/R. This simple integer state description will be used as an index for the residual shelf life of an RBC unit.

Phase-type distribution

The Erlang distribution, denoted by $\mathrm{Erl}\,(N, \nu_N)$, consists of N sequential phases, where the time spent in each phase is exponentially distributed with rate ν_N, as illustrated in Figure 5.5. It can be seen as a special case of the class of continuous phase-type distributions, which are typically used to model the time before a continuous-time Markov chain enters the absorbing state [80].

The continuous-time jump process $\tilde{\rho}$ is equivalent to an absorbing Markov chain with state space $\mathcal{R} = \{N, N-1, N-2, ..., 1, 0\}$, where states $N, N-1, ..., 1$ are transient states and state 0 is an absorbing state, and transition rate matrix

$$
Q = \begin{bmatrix}
-\nu_N & \nu_N & 0 & \cdots & 0 \\
0 & -\nu_N & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
\vdots & & \ddots & -\nu_N & \nu_N \\
0 & \cdots & \cdots & 0 & 0
\end{bmatrix}
= \left[\begin{array}{c|c} G & \boldsymbol{g}_0 \\ \hline \mathbf{0} & 0 \end{array}\right],
\tag{5.9}
$$

where $G \in \mathbb{R}^{N \times N}$, $\boldsymbol{g}_0 \in \mathbb{R}^{N \times 1}$, $\mathbf{0} \in \mathbb{R}^{1 \times N}$, and 0 is a scalar. Moreover, the initial state distribution is given by $[\boldsymbol{\alpha} \mid \alpha_0] = [1\ 0\ \cdots\ 0 \mid 0]$, where $\boldsymbol{\alpha} \in \mathbb{R}^{1 \times N}$ and α_0 is a scalar.

Definition 5.1 (Phase-type distribution of T_N)**.** The time until the random variable T_N enters the absorbing state has phase-type distribution $PH(\boldsymbol{\alpha}, G)$. The distribution function of T_N can be expressed as

$$
\begin{aligned}
F_{T_N}(t) &= 1 - \boldsymbol{\alpha} \exp(Gt)\boldsymbol{e}, \\
&= 1 - \exp(-\nu_N t) \sum_{k=0}^{N-1} \frac{(\nu_N t)^k}{k!},
\end{aligned}
\tag{5.10}
$$

for $t \geq 0$, where $\exp(Gt)$ is the matrix exponential, and $\boldsymbol{e}$ is the all ones vector. The moments of T_N are given by

$$
\begin{aligned}
\mathbb{E}\left(T_N^k\right) &= (-1)^k k! \boldsymbol{\alpha} G^{-k} \boldsymbol{e}, \\
&= \frac{(N+k-1)!}{(N-1)!} \nu_N^{-k},
\end{aligned}
\tag{5.11}
$$

implying that $\mathbb{E}\,(T_N) = N\nu_N^{-1}$ and $\mathrm{Var}\,(T_N) = N\nu_N^{-2}$.

As mentioned before, the main advantage of approximating the maximum shelf life of an RBC unit by a phase-type distribution (i.e., we keep track of the number of residual phases left before the RBC unit expires) is that the GSMP description of the inventory allocation problem can be transformed into continuous-time Markov

Figure 5.5: Phase diagram of an $\mathrm{Erl}(N, \nu_N)$ distribution.
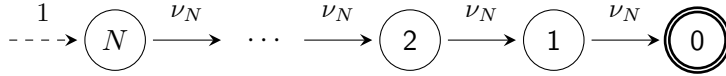
chain. This entire setup can also be expanded to multiple RBC units, as will be shown in the next section. Eventually, in Section 5.4, this will lead to a relatively simple formulation of the GSMDP as a discrete-time MDP with just a discrete state, a discrete action, and discrete time description and three types of transitions: donations, requests, and ageing.

Remark 5.2 (Ageing parameter)**.** In the remainder of this chapter we will use ν instead of ν_N to refer to the ageing parameter.

5.3.2 Multiple RBC units

The same approach as in Section 5.3.1 can also be applied to multiple RBC units. The original outdating process has the natural structure of GSMP , i.e., at arbitrary times elements can be triggered or enabled that exist for random periods of time. These period run parallel or synchronously but independent. For the current special case of blood inventory, however, these times incorporate:

- an exponential time for a next donation,

- an exponential time for a next request,

- deterministic times for experiation of the individual RBC units in inventory

This system can be modeled as a continuous time Markov process. One characterization of the state space is by complete account of the inventory. That is, the state variable would be

$$\left\{ r_1(t), ..., r_{L(t)}(t) \right\}, \tag{5.12}$$

where $L(t)$ is the number of units in inventory at time t and $r_k(t)$ is the residual shelf life of the kth RBC unit in inventory where the units are ranked by their residual shelf life, for $k = 1, ..., L(t)$. In this form, however, the analysis of the process is difficult, if not impossible, due to the varying dimensions of the state space. As an alternative, we might aggregate RBC units with the same residual shelf life, i.e., $\boldsymbol{s} = [s_1, ..., s_N]^\top$, where s_r indicates the number of RBC units in

Figure 5.6: Tandem queue consisting of N infinite server queues.
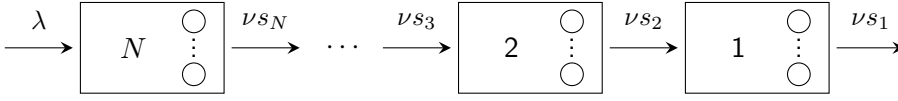
inventory with residual shelf life $r \in \mathcal{R}$ and transition rate matrix $Q \in \mathbb{R}^{\mathcal{S} \times \mathcal{S}}$, with

$$q\left(s' \mid s\right) = \begin{cases} \lambda, & \begin{cases} s'_N = s_N + 1 \\ s'_\ell = s_\ell & \ell < N \end{cases} \\ \nu s_r, & \begin{cases} s'_\ell = s_\ell - 1 & \ell = r \\ s'_\ell = s_\ell + 1 & \ell = r - 1 \quad \forall r > 1 \\ s'_\ell = s_\ell & \ell \neq r \end{cases} \\ \nu s_1 & \begin{cases} s'_1 = s_1 - 1 \\ s'_\ell = s_\ell & \ell > 1 \end{cases} \end{cases} \tag{5.13}$$

for $s' \neq s$ and $q(s'|s) = -\sum_{s' \in \mathcal{S}: s' \neq s} q(s'|s)$ if $s' = s$

Independent infinite-server tandem queues

A different way of viewing the Erlangian phase-type distribution is by regarding it as a tandem queue, with single server stations (for a single RBC unit) or with infinite servers stations (for multiple units). More detailed, for the multiple unit case, consider a tandem network of N queues that are placed in series. Each queue has an infinite number of servers with exponentially distributed service time, with mean ν. Fresh RBC units arrive according to a Poisson process with rate λ. Upon service completion at queue r the RBC unit routes to queue $r-1$, $r = N, ..., 2$, and finally departs from queue 1, as illustrated in Figure 5.6. For each residual shelf life $r \in \mathcal{R}$ the amount of RBC units in inventory behaves like an infinite-server tandem queue.

Steady state distribution

The major advantage of this representation is that we only need to keep track of the total number for different blood groups $i \in \mathcal{B}$ and phases $r \in \mathcal{R}$ and not the individual blood units themselves. This will become important in Section 5.4, where the RBC unit to issue will be selected. By the representation given here, and in line with the inventory, the following result can already be concluded from literature on so-called product forms (e.g., [116]) if no demands would take place.

$$\mathbb{P}\left[S_N = s_N, ..., S_1 = s_1\right] = \mathbb{P}\left[S = s\right] = \exp\left(-\frac{N\lambda}{\nu}\right) \prod_{r=1}^{N} \frac{1}{s_r!} \left(\frac{\lambda}{\nu}\right)^{s_r}. \tag{5.14}$$

This theoretical result could be of interest to compare with an actual distribution of RBC units in inventory as due to blood units meanwhile been issued, with and without compatibility taken into account

5.4 Markov decision process for blood issuing

Now that the structure of the underlying blood dynamics has been brought down to a GSMP and even, due to the exponential age modelling, a continuous time Markov process, more precisely chain as the states are discrete, CTMC, the optimization of issuing actions can be incorporated. This involves two additional main steps:

- to adopt and apply the theory of Markov decision processes, generally referred to as MDP,

- to transform a continuous-time MDP into a discrete-time MDP

So far, no requests for RBC units nor issuing decisions have been considered. Therefore, this section describes the blood inventory management process as a Markov decision process (MDP). This is a mathematical framework for sequential decision making in a stochastic environment based on the Markovian assumption.

For the inventory allocation problem, it implies that the MDP model determines which RBC unit should be selected from inventory to satisfy a request for a RBC unit with a particular blood group, such that

- the selected RBC unit is compatible with the RBC unit requested,

- the probability that future requests can be satisfied is maximized,

- the probability that units get outdated is minimized.

5.4.1 Continuous-time MDP

An infinite horizon continuous time MDP is generally defined by a quin-tuple $(\mathcal{T}, \mathcal{S}, \mathcal{X}, Q, c)$, corresponding with decision epochs, state space, action space, transition matrix, and cost matrix respectively.

Decision epochs

The set of time epochs is given by $\mathcal{T} = [0, \infty)$, where an epoch $t \in \mathcal{T}$ represents either the arrival of a fresh RBC unit (at a rate λ), the ageing of an RBC unit (at a rate ν), or the a request for a single RBC unit (at a rate μ) each of which can take place any time.

State space

The state of the inventory can mathematically be described by a vector

$$\boldsymbol{s} = \begin{bmatrix} \underbrace{s_{11} \quad \cdots \quad s_{1N}}_{N} & \underbrace{s_{21} \quad \cdots \quad s_{2N}}_{N} & \cdots & \underbrace{s_{m1} \quad \cdots \quad s_{mN}}_{N} \end{bmatrix}^{\top} \qquad (5.15)$$

where s_{ir} indicates the number of RBC units in inventory with blood group $i \in \mathcal{B}$ and residual shelf life $r \in \mathcal{R}$. To limit the size of the state space, the maximum number of RBC units in inventory can be bounded by a constant S.

Action space

Although three different types of events can occur (see Section 5.2.1), only one of them is associated with an action. Upon arrival of a request for RBC units with a particular blood group, a double decision is to be made: the decision maker either selects compatible RBC units from inventory to satisfy this request or leaves it unsatisfied. A request can be left unsatisfied if their are either no compatible RBC units present in inventory or the optimal action is not to satisfy the request, even when there are compatible units available. In both cases, we will refer to such an event as a shortage. Given that blood group $j \in \mathcal{B}$ is requested the possible actions are:

$$\mathcal{X}_j(\boldsymbol{s}) = \{(i,r) \mid s_{ir} > 0, \; i \leq j\} \cup \{u\}, \qquad (5.16)$$

with

$$|\mathcal{X}_j(\boldsymbol{s})| = \sum_{i \in \mathcal{B} \; \mid \; i \leq j} \sum_{r \in \mathcal{R}} s_{ir} + 1. \qquad (5.17)$$

Moreover, for each blood group $j \in \mathcal{B}$ an action has to be chosen from the set of feasible actions, given by

$$\mathcal{X}(\boldsymbol{s}) = \{\boldsymbol{x} \in \mathbb{R}^m \mid x_j \in \mathcal{X}_j(\boldsymbol{s})\}. \qquad (5.18)$$

Transition rates

Assume that fresh RBC units arrive at the inventory according to a Poisson process at rate λ and that requests occur according to an independent Poisson process at rate μ. Moreover, assume that the probability that a donated RBC unit has blood group $i \in \mathcal{B}$ and the probability that a requested RBC unit has blood group $j \in \mathcal{B}$ is consistent with the distribution of blood groups in the Caucasian population. Hence,

$$\lambda_i = \lambda \cdot f_i \qquad \text{and} \qquad \mu_j = \mu \cdot f_j \qquad (5.19)$$

Figure 5.7: Tandem queue consisting of N infinite server queues.
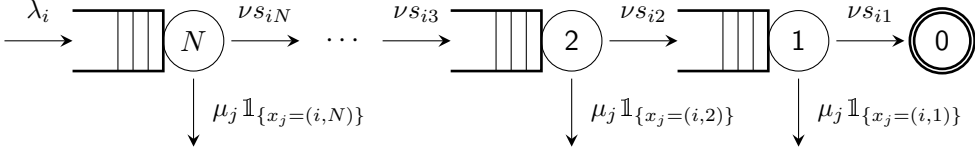
The transition rate matrices $Q \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}| \times |\mathcal{X}|}$ for an arbitrary action $x \in \mathcal{X}$ can be defined as

$$
q\left(s' \mid s, x\right) = \begin{cases}
\lambda_i \mathbb{1}_{\{|s| \leq S\}}, & \begin{cases} s'_{kN} = s_{kN} + 1 & k = i \\ s'_{kN} = s_{kN} & k \neq i \\ s'_{k\ell} = s_{k\ell} & \ell < N \end{cases} & \forall i \in \mathcal{B}, \\[2em]
\mu_j \mathbb{1}_{\{x_j=(i,r)\}} & \begin{cases} s'_{k\ell} = s_{k\ell} - 1 & k = i,\ \ell = r \\ s'_{k\ell} = s_{k\ell} & k \neq i,\ \ell \neq r \end{cases} & \forall j \in \mathcal{B}, \\[2em]
\nu s_{ir}, & \begin{cases} s'_{k\ell} = s_{k\ell} - 1 & k = i,\ \ell = r & \forall i \in \mathcal{B} \\ s'_{k\ell-1} = s_{k\ell-1} + 1 & k = i,\ \ell = r & \forall r \in \mathcal{R} \\ s'_{k\ell} = s_{k\ell} & k \neq i,\ \ell \neq r & r > 1, \end{cases} & \\[2em]
\nu s_{i1} & \begin{cases} s'_{k1} = s_{k1} - 1 & k = i \\ s'_{k1} = s_{k1} & k \neq i \\ s'_{k\ell} = s_{k\ell} & \ell > 1 \end{cases} & \forall i \in \mathcal{B},
\end{cases}
\tag{5.20}
$$

for a transition from state $s \in \mathcal{S}$ into another state $s' \neq s$, with $s' \in \mathcal{S}$, under decision $x \in \mathcal{X}(s)$, and for the diagonal elements $s' = s$:

$$
\begin{aligned}
q\left(s' \mid s, x\right) &= -\sum_{\substack{s' \in \mathcal{S} \\ s' \neq s}} q\left(s' \mid s, x\right) \\
&= -\left(\lambda \mathbb{1}_{\{|s| \leq S\}} + \mu + \nu |s|\right) + \underbrace{\sum_{j \in \mathcal{B}} \mu_j \mathbb{1}_{\{x_j = u\}}}_{\text{refused request}}.
\end{aligned}
\tag{5.21}
$$

Here it is indicated that the last term in equation (5.21), which corresponds to the action that a request is refused, as will be relevant for cost inclusion.

Costs

A second aspect for the CTMC description to be expanded to a MDP is to include a cost or reward structure. For the blood inventory problem three different types of costs can be thought of:

- c_{ijr}: costs for allocation an RBC unit with blood group $i \in \mathcal{B}$ and residual shelf life $r \in \mathcal{R}$ to a request for an RBC unit with blood group j.

- c_s: shortage costs if a request is refused.

- c_o: outdating costs if an RBC unit expires.

Due to the additive structure for an expected cumulative reward measure, these can be aggregated as the costs per unit of time in a given state and action [89]:

$$c\left(\boldsymbol{s}, \boldsymbol{x}\right) = \sum_{i \in \mathcal{B}} \sum_{j \in \mathcal{B}} \sum_{r \in \mathcal{R}} c_{ijr} \mu_j \mathbb{1}_{\{x_j = (i,r)\}} + c_s \sum_{j \in \mathcal{B}} \mu_j \mathbb{1}_{\{x_j = u\}} + c_o \sum_{i \in \mathcal{B}} \nu s_{i1}. \quad (5.22)$$

5.4.2 Uniformization

Now that a discrete and relatively simple state transition structure has been established, a computational scheme is required to determine 'optimal' actions for any state $s \in \mathcal{S}$. Here different cost aspects can be incorporated as specified in equation (5.22). The general objective is to minimize average costs, i.e., as the average over a long period of time, expressed by

$$G(\boldsymbol{s}) = \min_{\delta \in \Delta} \lim_{Z \to \infty} \frac{1}{Z} \int_0^Z \mathbb{E}\left[c\left(S_\tau, \delta\left(S_\tau\right)\right) \mid S_0 = \boldsymbol{s}\right] d\tau \quad (5.23)$$

where δ represents a decision rule, which prescribes a decision vector $\boldsymbol{x} \in \mathcal{X}$ for any possible state $\boldsymbol{s} \in \mathcal{S}$, and $\Delta = \{\delta : \mathcal{S} \to \mathcal{X}\}$ is a set of decision rules. Note that the cost function c itself already includes an expectation with respect to possible blood groups requested, as specified in equation (5.22). Furthermore, the average expected values $G(\boldsymbol{s})$, conditional on the initial state $\boldsymbol{s}$, can be assumed to be independent of the initial state under general ergodic conditions for the continuous-time Markov chains and decision rule δ (e.g. see [89]). $G(s)$ can therefore be defined as a scalar

$$g^* = G(\boldsymbol{s}) \qquad \forall s \in \mathcal{S}. \quad (5.24)$$

In addition, as also justified in [89], a stationary decision rule, (i.e., independent of time) can be shown to be average optimal, among all possible strategies that in principle allow a different rule at any time.

Despite the discrete state description, from a computational point of view, the optimal value g^* still involves the computation of the minimum of a limit integral in continuous time. This will remain highly complex, if not impossible. As a final essentially simplifying step, the approach of uniformization (or randomization) will be brought in. Intuitively, this approach can be regarded as a time discretization. More precisely, let B be an arbitrarily large finite number such that

$$\lambda + \mu + \nu S \leq B < \infty. \quad (5.25)$$

We can then transform the continuous-time MDP into a more computable discrete-time MDP by regarding small time steps of length $h = 1/B$. We refer to remark 5.3 below for more details on this approach and related literature. Define the discrete-time MDP by the one-step transition probabilities

$$p\left(\boldsymbol{s}' \mid \boldsymbol{s}, \boldsymbol{x}\right) = \begin{cases} \frac{\lambda_i}{B} \mathbb{1}_{\{|s| \leq S\}} & \begin{cases} s'_{kN} = s_{kN} + 1 & k = i \\ s'_{kN} = s_{kN} & k \neq i \\ s'_{k\ell} = s_{k\ell} & \ell < N \end{cases} & \forall i \in \mathcal{B}, \\[20pt] \frac{\mu_j}{B} \mathbb{1}_{\{x_j = (i,r)\}} & \begin{cases} s'_{k\ell} = s_{k\ell} - 1 & k = i, \ \ell = r \\ s'_{k\ell} = s_{k\ell} & k \neq i, \ \ell \neq r \end{cases} & \forall j \in \mathcal{B}, \\[20pt] \frac{\nu}{B} s_{ir} & \begin{cases} s'_{k\ell} = s_{k\ell} - 1 & k = i, \ \ell = r & \forall i \in \mathcal{B} \\ s'_{k\ell-1} = s_{k\ell-1} + 1 & k = i, \ \ell = r & \forall r \in \mathcal{R} \\ s'_{k\ell} = s_{k\ell} & k \neq i, \ \ell \neq r & r > 1, \end{cases} \\[20pt] \frac{\nu}{B} s_{i1} & \begin{cases} s'_{k1} = s_{k1} - 1 & k = i \\ s'_{k1} = s_{k1} & k \neq i \\ s'_{k\ell} = s_{k\ell} & \ell > 1 \end{cases} & \forall i \in \mathcal{B}, \\[20pt] 1 - \frac{1}{B} \left(\lambda \mathbb{1}_{\{|s| \leq S\}} + \sum_{j \in \mathcal{B}} \mu_j \mathbb{1}_{\{x_j \neq u\}} + \nu |s| \right) & \boldsymbol{s}' = \boldsymbol{s} \end{cases}$$

$$(5.26)$$

Remark 5.3 (Uniformization). For a more extensive and intuitive insight in the approach of uniformization, the interested reader is referred to [118]. Like in this reference, equality of a continuous-time and uniformized discrete-time Markov chain (or Markov Process) is intuitively obvious for an average cost case or performance measure. As proven in [72, 105], for the average case it can be extended to also include actions to compute an average optimal strategy. As in [118], a uniformization equivalence can be shown for any finite time period and any given set of actions. This, however, would need a Poissonian expansion. In addition, time dependent strategies could then be expected. In order to justify a more simple formula for approximate iteration, similar to equation (5.29) in the next section, a discrepancy arises. An error bound on the accuracy can then be proven to be linear in h [115]. As a result, it vanishes if h is taken arbitrarily small [115]. In this chapter we just consider the average cost case.

5.4.3 Stochastic Dynamic Programming

Now that the original blood inventory allocation has been fully taken down to a discrete-state as well as discrete-time MDP, the standard procedures for MDP modelling and computation, as extensively described in [89], can be applied. More

precisely, based on uniformization and MDP the average optimal value can be presented as

$$g^* = \lim_{N\to\infty} \frac{1}{Nh} \sum_{k=0}^{N-1} \left[P^{\delta^*}\right]^k c(s, \delta^*(s)) \qquad \text{for any initial state } s \in \mathcal{S}. \tag{5.27}$$

where

δ^*:   the decision rule with optimal actions $x^* = \delta^*(s)$ in any state $s \in \mathcal{S}$,

P^{δ^*}:   the one-step transition probability matrix $p(s' \mid s, x)$ as by (5.26) with $x = \delta^*(s)$.

The optimal value g^*, the optimal decision rule δ^*, and its corresponding optimal actions x^* are implicitly determined by the following equation:

$$
\begin{aligned}
H(s) &= \min_{x \in \mathcal{X}(s)} \left\{ [c(s,x) - g^*] \cdot h + \sum_{s' \in \mathcal{S}} p(s' \mid s, x) H(s') \right\} \\
&= [c(s, \delta^*(s)) - g^*] \cdot h + \sum_{s' \in \mathcal{S}} p(s' \mid s, \delta^*(s)) H(s')
\end{aligned}
\tag{5.28}
$$

for all states $s \in \mathcal{S}$, where $H(s)$ is a function still to be determined. In words, δ^* determines an action in which the minimum value in equation (5.28) is achieved. Here the function H is known as the so-called relative gain function (see for example [89] and [13] Chapter 1). Particularly, for the study of the function H in depth the interested reader is referred to [111]. However, as (5.28) still involves an implicit equation, either of two iterative approaches are generally followed:

- policy improvement,

- successive approximation.

In fact, as the first is not necessarily guaranteed to always work and generally more complicated, the second approach is chosen and is briefly presented below. Here, for interpretation,

$V_k(s)$ can be seen as the optimal cumulative expected costs over k steps (each of time length h) when starting in state s at time 0.

First, let us start by setting $V_0(s) = 0$ for all $s \in \mathcal{S}$. Next, we iterate according to the successive approximation and the following backward recursive equation:

$$V_{k+1}(s) = \min_{x \in \mathcal{X}(s)} \left\{ h \cdot c(s,x) + \sum_{s' \in \mathcal{S}} p(s' \mid s, x) V_k(s') \right\}, \tag{5.29}$$

Note that for different values of k different optimizing actions might appear in one-and-the-same state, it is at least intuitively appealing that

$$\lim_{k\to\infty} \frac{1}{k} V_k(\boldsymbol{s}) \to h g^*$$ (5.30)

independent of $\boldsymbol{s}$. In fact, the so-called 'Odoni bounds' can be used to capture the value g^* up to some inaccuracy (see [82, 111]). Assuming that the cost function is non-negative, let

$$\underline{g}_k^* = \min_{s\in\mathcal{S}} \left\{ V_{k+1}(\boldsymbol{s}) - V_k(\boldsymbol{s}) \right\} / h$$

$$\overline{g}_k^* = \max_{s\in\mathcal{S}} \left\{ V_{k+1}(\boldsymbol{s}) - V_k(\boldsymbol{s}) \right\} / h$$

where $\underline{g}_k$ and $\overline{g}_k$ are the minimum and maximum difference of the value function over two iterations steps, k and $k+1$ respectively. Then it can be shown that

$$\underline{g}_k^* \leq \underline{g}_{k+1}^* \leq g^* \leq \overline{g}_{k+1}^* \leq \overline{g}_k^*$$ (5.31)

Furthermore, after a finite number of steps the optimizing actions will generally appear to remain unchanged. In line of the approach of policy improvement, these actions can then be expected to be optimal, or at least ϵ-optimal (up to a small inaccuracy ϵ) to obtain g^*.

5.5 Evaluation

In this chapter, a microscopic computational structure has been established for RBC issuing. It incorporates the realistic aspects of donations and requests at arbitrary times, as well as ageing of RBC units and blood group compatibility. For realistically sized situations this chapter and its formulations provides a base description for other approximative techniques (i.e, approximate dynamic programming or reinforcement learning), which are based upon MDP formulations.

5.6 Appendix I

5.6.1 Markov processes

The most standard Markov process is a discrete-time Markov chain, where the times between successive events are deterministic, implying that state transitions occur at unit time moments (i.e., $k = 1, 2, 3, ...$). In practice, events may occur at any point in time. In the continuous-time analogue of discrete-time Markov chains the times between successive state transitions are not deterministic, but exponentially distributed. For more properties of Markov processes we refer the reader to, for example, [97, 110].

Definition 5.2 (Markov process). The stochastic process $X = \{X(t) \mid t \geq 0\}$ with discrete state space $\mathcal{S}$ is called a Markov process if for each $t \geq 0$,

$$\mathbb{P}\left[X(t_{k+1}) = s(t_{k+1}) \mid X(t_0) = s(0), ..., X(t_k) = s(k)\right]$$
$$= \mathbb{P}\left[X(t_{k+1}) = s(k+1) \mid X(t_k) = s(k)\right],$$

for all $0 \leq (t_0) < ... < t_{k+1}$ and $s(0), ..., s(k+1) \in \mathcal{S}$.

   If the time between two state transitions follows an arbitrary continuous distribution, so it is no longer deterministic or exponentially distributed, then it is called a semi-Markov process (SMP).

5.6.2 Generalized semi-Markov process

As mentioned in Section 5.2, the dynamics between the supply and demand of RBC unit possesses, to a large extent follow the structure of a generalized semi-Markov (decision) process, which preserves the typical structure of a continuous-time stochastic discrete event system (with decisions) [49, 91, 131]. The abstract description of a GSMP given below is based on these three papers.

   A time-homogeneous GSMP consists of a countable set of states $\mathcal{S}$ and a countable set of events $\mathcal{E}$. At any point in time, the system occupies a state $s \in \mathcal{S}$ in which a subset $\mathcal{E}(s)$ of the events are enabled. Each event $e \in \mathcal{E}(s)$ is associated with a positive distribution G_e –governing the time until event e triggers if it continuously enabled– and a next-state probability distribution $p_e(s' \mid s)$. The events represent concurrent processes competing for the next state transition. To each enabled event $e \in \mathcal{E}(s)$, we associate a clock c_e reading the remaining time until event e occurs. Let e^* be the event with the smallest clock value, i.e.,

$$e^* = \arg\min_{e \in \mathcal{E}(s)} \{c_e\}.$$

When e^* triggers, the next state $s' \in \mathcal{S}$ is sampled according to the probability distribution $p_{e^*}(s' \mid s)$ and the clocks are updated:

$$c'_e = \begin{cases} c_e - c_{e^*} & \text{if } e \in \mathcal{E}_{s'} \text{ and } e \in \mathcal{E}_s \setminus \{e^*\} \\ \text{sampled from } G_e & \text{if } e \in \mathcal{E}_{s'} \text{ and } e \notin \mathcal{E}_s \setminus \{e^*\} \\ \infty & \text{if } e \notin \mathcal{E}_{s'} \end{cases}$$

In other words, events that remain enabled without triggering are not rescheduled, while the triggering event and newly enabled events are. This procedure is repeated with $s = s'$ and $c_e = c'_e$.

For a simulation approach this description would be appropriate. For exact and numerical computations, however, general and in particular deterministic asynchronous clocks are most hard to deal with. In Section 5.3, these timers will therefore be approximated by so-called phase-type distributions, more precisely Erlang distributions. This will transform the natural description for blood units, as in Section 5.2 into a more tractable continuous-time Markov process. In particular, the residual shelf life for RBC units in the inventory can be traced in a countable manner.

Optimal blood issuing by comprehensive matching

6.1 Introduction

6.1.1 Blood Management

Hospitals in the Netherlands use approximately 400,000 red blood cell (RBC) units each year [99]. These units are donated by voluntary non-remunerated donors at collection sites throughout the country. After collection, the donated blood units are processed (i.e., separated in red blood cells, platelets, and plasma), typed for more than fourteen different antigens, tested for several infectious diseases and finally stored in one of the distribution centers until requested by hospitals. Although describing the process in this way makes it sound like a straightforward process, several complications arise, making inventory management of RBC units an important and interesting topic for research.

First of all, RBC units are used during major surgeries or as treatment for leukemia, anemia, and blood disorders. Not being able to satisfy requests from hospitals comes at a very high cost, since this may lead to delays and therefore places transfusion recipients at risk. Hence, adequate and timely availability of RBC units is essential.

Second, RBC units are perishable products. After 35 days of storage, the unit has to be discarded. In the inventory management for perishable products, a balance has to be found between the probability of outdating and the probability of shortage, as decreasing one usually increases the other. Moreover, RBC units are obtained from donations by voluntary, non-remunerated donors. Donors are motivated by the fact that their blood donation is necessary and saves lives. Increasing outdating would affect their motivation and is therefore undesirable. So, for RBC units and other blood products, both outdating and shortages should be minimized at the least, and preferably prevented altogether.

Third, differences in blood groups pose another challenge. The ABO blood

groups are well known, but many more blood groups exist. Although compatible issuing can be done, identical issuing is to be preferred, as compatible issuing may cause shortages for RBC units with a rare blood group.

Fourth, inventory management of RBC units has to deal with stochasticity. Most inventory management systems deal with stochastic demand. However, the supply side of blood inventory also contains stochasticity. As donors are donating voluntarily and are non-remunerated, the probability of a no-show for a donation is substantial. Additionally, as Sanquin does not use appointments for blood donors, the lead-time between inviting a donor and the donor showing up is highly uncertain, ranging from a week up to a month.

These four complications (i.e., shortage, outdating, blood group compatibility, and stochasticity in the supply and demand of RBC units) make inventory management of RBC units non-standard and non-trivial. This chapter presents a mathematical framework for the inventory management of RBC units. Besides minimizing outdating and shortages, we will also take the rareness of a blood group into account.

6.1.2   Approach

To investigate an issuing policy that finds a trade-off between the age and rareness of the RBC units in inventory, a mathematical framework will be presented, which consists of three major components:

i) a binary representation of general blood groups (see Chapter 4),

ii) a deterministic minimal cost flow problem (MCFP), which finds an optimal allocation between the RBC units in inventory and RBC units requested, while taking all haematological constraints with respect to blood group matching into account,

iii) a simulation model, which iteratively solves the MCFP, incorporates the stochasticity in the supply and demand of RBC units, and evaluates the long-term performance of a proposed issuing policy (e.g., percentage of RBC units short, percentage of RBC units outdated, percentage of blood group identical matches, average issuing age).

As opposed to literature, the mathematical framework presented in this chapter is novel in that it can be applied for general blood groups (previously only the ABO, D blood groups were included) and models the inventory allocation problem as an MCFP. In addition, the concept of 'relative opportunity loss' is introduced to save RBC units with rare blood groups.

6.1.3 Outline

This chapter is organized as follows. Section 6.2 provides an overview of the relevant literature on inventory management of blood products. In Section 6.3 we show 1) that the inventory allocation problem can be described as a transportation problem, and 2) that the objective function of the MCFP can be used to represent an issuing policy that balances the rareness and age of the RBC units in inventory. Next, in Section 6.4, a simulation model is presented to incorporate the stochasticity in the supply and demand of RBC units and evaluate the long-term performance of the proposed issuing policy. The results of the computational experiments are discussed in Section 6.5. Finally, Section 6.6 concludes the chapter and indicates directions for future research.

6.2 Literature

Inventory management of perishable products is an important topic in Operations Research. [4, 50, 62, 79] gave a review of the available literature with respect to this topic. A part of this literature consists of the inventory management of blood products. The management of these specific products also garners much attention itself. [7, 83, 86] are three recent review papers. [7] classified the papers according to the blood component under study: platelets, RBCs, plasma, whole blood, frozen blood, other/unclear. Whereas we are interested in the inventory management of perishable products we focus on the literature about platelets and RBCs. The most important difference with respect to the inventory management of these two blood components is their maximum shelf life. Platelets are considered to be expired after 7 days and RBCs after 35 days.

In Section 6.2.1 we will review the recent literature on the inventory management of platelets, and continue afterwards with the literature on the inventory management of RBC in Section 6.2.2. Finally, in Section 6.2.3 we will discuss where our research builds upon current approaches and extends existing literature.

6.2.1 Platelets

Whereas platelets expire after 7 days most papers about platelet inventory management consider the percentage of outdated units as main performance measure and apply a First-In-First-Out (FIFO) issuing policy [41, 77, 106]. Though, the inventory size should be sufficiently large to prevented or control shortages. One way to control shortages is by using a predetermined maximum shortage level [30]. Another way to control both outdating and shortages is by including both performance measures into the objective function [1, 23].

A FIFO issuing policy seemed to be optimal for inventory management of platelets. However, as for some patient groups fresh platelets increase survival

rates [54, 117], some papers use different issuing policies. One way to incoporate this is by making a difference between 'young' and 'old/any' platelets [22, 52]. Moreover, whereas [52] looked at the inventory management of hospitals they incorporated a cross-match-to-transfusion ratio and the cross-match release period in their models. [23] shortened the maximum shelf life of the platelets to three days and classified them as 'young', 'mature', and 'old'. Moreover, they included protection levels and substitution costs to limit the amount of 'young' platelets issued to satisfy requests where the age of the platelets did not matter.

With respect to replenishment policies [30] considered an old inventory ratio policy to avoid shortages. This policy states that if the proportion of old units in stock exceeded a certain threshold, making it likely that some units will get outdated, some extra donors should be invited. A similarly approach was proposed by [54, 55] and [28]. They investigated 1D and 2D order-up-to-rules, where in the 2D order-up-to-rule donors where invited based on both the total amount of platelet units in stock and the amount of young platelet units in stock. Recently, [92] analyzed several ordering policies for platelets which also consider hospital size and demand variation.

6.2.2   Red blood cells

For the inventory management of RBCs the percentage of outdating and shortage remain important performance measures as well. Though, some papers suggest that the maximum shelf life of 35 days may be reduced without excessive increases of outdating and shortage rates [10, 31].

[29] developed a two stage inventory control model, where in the first stage decisions about the review period and order-up-to-levels are made and in the second a stage decision about the daily operation of the system are considered. Moreover, they investigate the difference between exact and compatible issuing policies with respect to ABO, RhD compatibility. A similar modelling approach is applied by [56].

[3] apply a single threshold policy in which blood younger than the threshold is issued according to a FIFO policy and blood that is older is issued according to a LIFO policy. For a threshold of 14 days they show that the mean age of transfused blood decreases from 30 to 20 days.

6.2.3   Relevance of this work

Within the existing literature on the inventory management of blood products no one has yet considered more than the common eight ABO, D blood groups. However, since the beginning of the 21th century, hospitals are recommended to extend their matching strategies for some patient categories [18]. As a result, transfusion recipients belonging to such patient categories are not only matched

for the standard ABO, D blood groups, but also for some additional blood groups (e.g., Kell, Duffy, Kidd, MNS). These extended matching strategies lead to a growing demand for more specific blood groups.

Clearly, by including more antigens both the diversity among the blood units in inventory and the diversity among the blood units requested increases. Hence, the likelihood of finding an exact match between a unit requested and a unit issued decreases. Therefore, we extend the existing performing measures such as outdating, shortage, and issuing age by a fourth performance measure: the relative opportunity loss of a match.

In the primary objective of the (mathematical) model that will be presented in this chapter we aim to balance the relative opportunity loss of a match and the age of an issued RBC unit (FIFO). Where minimizing the relative opportunity loss of a match will correspond to saving rare units to prevent shortages and applying a FIFO policy corresponds to the prevention of outdating, these performance measures are not explicitly included in the objective of the mathematical model. However, the simulation that evaluates the quality of a match and age of issued units over time, also keeps track of these performance measures.

6.3   Inventory allocation problem

In the classical transportation problem products have to be transported from a set of sources $\mathcal{V}_1$ to a set of destinations $\mathcal{V}_2$ at minimum costs, such that all supply and demand constraints are satisfied [9, 128]. Its applicability to blood inventory allocation seems straightforward, as RBC units have to be transported from distribution centers to hospitals. However, it is not the transportation issue itself that is of interest, but the decisions about which RBC units to select from inventory, such that

- all requests can be satisfied with compatible RBC units from inventory,

- shortages for future requests are avoided,

- outdating is prevented (RBC units are perishable),

as illustrated in Figure 6.1.

To show which steps are required to transform the inventory allocation problem into a transportation problem, Section 6.3.1 first illustrates them for the well-known ABO blood groups. In Sections 6.3.2 and 6.3.3 the same steps are described in a unified mathematical formulation, such that the inventory allocation problem can be solved for general blood groups.
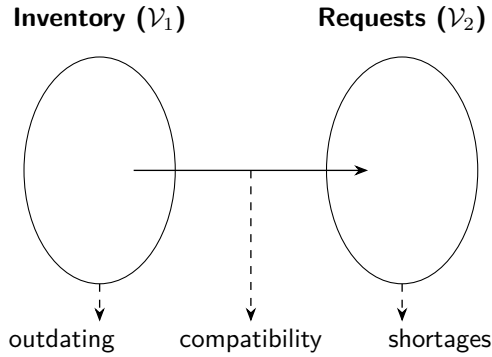
Figure 6.1: Schematic representation of the inventory allocation problem, where $\mathcal{V}_1$ represents the set of RBC units in inventory and $\mathcal{V}_2$ represents the set of RBC units requested. The aim is to find an antigen compatible match between the RBC units in inventory and the RBC units requested, such that both shortages and outdating are minimized.

6.3.1   Illustrative ABO example

Consider an example of the inventory allocation problem as presented in Figure 6.2, where $\mathcal{V}_1$ represents the set of RBC units supplied and $\mathcal{V}_2$ represents the set of RBC units requested. A vertex $(i, r) \in \mathcal{V}_1$, with

i: blood group supplied,
r: residual shelf life,

refers to all RBC units in inventory with blood group i and a residual shelf life of r days. A vertex $j \in \mathcal{V}_2$, with

j: blood group requested,

refers to all RBC units requested with blood group j. Two vertices $(i, r) \in \mathcal{V}_1$ and $j \in \mathcal{V}_2$ are connected by an edge $[(i, r), j]$, if blood group i is compatible with blood group j. The number of RBC units supplied by a vertex $(i, r) \in \mathcal{V}_1$ and the number of RBC units requested by a vertex $j \in \mathcal{V}_2$ are given by the parameters s_{ir} and d_j respectively.

The objective function of the inventory allocation problem aims to control the inventory of RBC units, such that shortages for future requests are avoided and outdating is prevented. Shortages can be avoided by saving RBC units with a rare blood group. Outdating can be prevented by issuing RBC units with a short residual shelf life. We therefore take the following objective function

$$c_{[(i,r),j]} = \gamma \cdot \Delta_{ij} + (1 - \gamma) \cdot \frac{r}{2}, \tag{6.1}$$
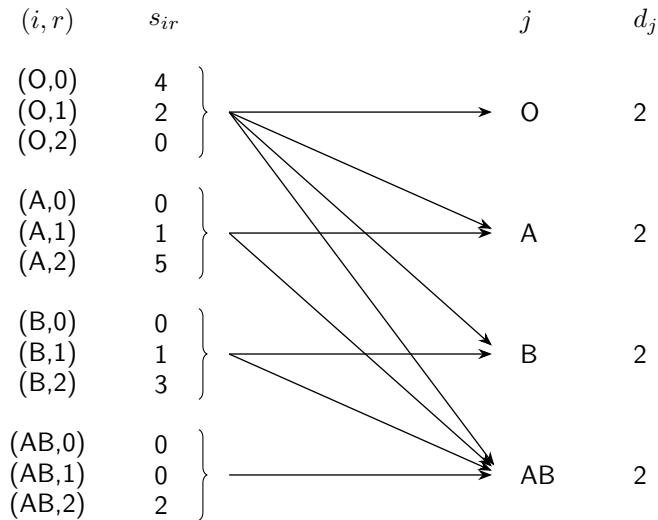
Figure 6.2: ABO-example of the inventory allocation problem, with s_{ir} the number of RBC units in inventory with blood group $i \in \mathcal{B}$ and a residual shelf life of r days and d_j the number of RBC units requested with blood group $j \in \mathcal{B}$.

where $\gamma \in [0,1]$ is parameter balancing relative opportunity loss $\Delta_{ij} \in [0,1]$ with the residual shelf life of the RBC units in inventory $\frac{r}{2} \in [0,1]$. To illustrate how the choice of γ influences the optimal allocation, consider the following three values for γ:

- $\gamma = 0$: the issuing policy minimizes the residual shelf life of the RBC units in inventory $\left(c_{[(i,r),j]} = \frac{r}{2}\right)$, a First-In-First-Out (FIFO) issuing policy is applied,

- $\gamma = 1$: the issuing policy minimizes the relative opportunity loss (MROL) of a match $\left(c_{[(i,r),j]} = \Delta_{ij}\right)$,

- $\gamma = \frac{1}{2}$: the issuing policy balances the relative opportunity loss of a match and the residual shelf life of the RBC units in inventory, which are weight equally.

For the FIFO strategy $(\gamma = 0)$ all units with blood group O are issued. It is therefore likely that future request cannot be satisfied. For the MROL strategy $(\gamma = 1)$, where the focus is purely on minimizing the relative opportunity loss, two units with blood group O get outdated. For the combined strategy $(\gamma = \frac{1}{2})$ no units get outdated while also 2 units with blood group O remain in the inventory. Hence, the trade-off between the quality of a match and the age of the RBC units in inventory seems beneficial (Table 6.1).

Table 6.1: Inventory of the RBC units that remains after an optimal allocation of RBC have been determined for $\gamma = 0, \frac{1}{2}, 1$.

Issuing policy	O			A			B			AB		
	4	2	0	0	1	5	0	1	3	0	0	2
FIFO ($\gamma = 0$)	0	0	0	0	0	5	0	0	3	0	0	2
FIFO/MROL ($\gamma = \frac{1}{2}$)	0	2	0	0	0	5	0	0	3	0	0	0
MROL ($\gamma = 1$)	2	2	0	0	0	4	0	0	2	0	0	0

6.3.2 Mathematical optimization model

Now that the appropriate abstractions of haematological concepts have been made, the inventory allocation problem can be transformed into a classical transportation problem, as already illustrated in Section 6.3.1 for the ABO blood groups and described more generally in this section. The power of this transformation is that all complex haematological concepts are dissolved in the definitions of the sets and parameters of the transportation problem. As a result, large inventory allocation problems consisting of 2^{14} different blood groups and expiration dates up to 35 days remain solvable.

Inventory allocation problem

Let $\mathcal{G}(\mathcal{V}, \mathcal{T})$ be a directed bipartite graph with vertex set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$, and edge set $\mathcal{T}$, such that

- $\mathcal{V}_1 = \{(i, r) \mid i \in \mathcal{B}, \ r \in \mathcal{R}\}$, where a vertex $(i, r) \in \mathcal{V}_1$ represents all RBC units in inventory with blood group $i \in \mathcal{B}$ and residual shelf life of $r \in \mathcal{R}$,

- $\mathcal{V}_2 = \{j \mid j \in \mathcal{B}\}$, where a vertex $j \in \mathcal{V}_2$ represents all RBC units requested with blood group $j \in \mathcal{B}$,

- $\mathcal{T} = \{\tau = [(i, r), j] \mid i \leq j, \ (i, r) \in \mathcal{V}_1, \ j \in \mathcal{V}_2\}$, where an edge $\tau \in \mathcal{T}$ indicates that vertex $(i, r) \in \mathcal{V}_1$ is connected with vertex $j \in \mathcal{V}_2$, or stated differently, blood group $i \in \mathcal{B}$ is compatible with blood group $j \in \mathcal{B}$,

as illustrated in Figure 6.3. Note that the connection between two vertices $(i, r) \in \mathcal{V}_1$ and $j \in \mathcal{V}_2$ only depends on the compatibility between the blood groups i and j, and is independent of the residual shelf life of the issued RBC units. The vertices $(i, r) \in \mathcal{V}_1$ can thus be clustered based upon their common blood group i.

Each vertex $(i, r) \in \mathcal{V}_1$ has a non-negative supply $s_{ir} \in \mathbb{N}_0$, which indicates the number of RBC units in inventory with blood group $i \in \mathcal{B}$ and residual shelf life $r \in \mathcal{R}$. Similarly, each vertex $j \in \mathcal{V}_2$ has a non-negative demand $d_j \in \mathbb{N}_0$, which indicates the number of RBC units requested with blood group $j \in \mathcal{B}$. The
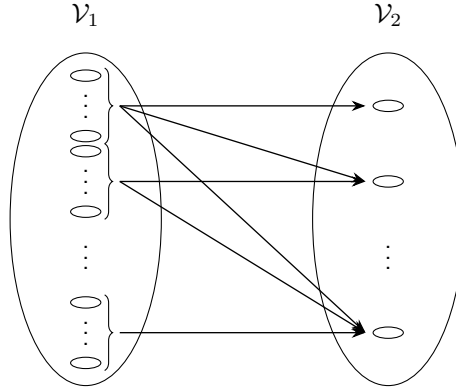
Figure 6.3: Graph $\mathcal{G}(\mathcal{V}, \mathcal{T})$ underlying the transportation problem, with vertex set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ and edge set $\mathcal{T}$. A vertex $(i, r) \in \mathcal{V}_1$ represents all RBC units in inventory with blood group $i \in \mathcal{B}$ and residual shelf life of $r \in \mathcal{R}$. A vertex $j \in \mathcal{V}_2$ represents all RBC units requested with blood group $j \in \mathcal{B}$.

maximum number of RBC units that can be allocated to an edge $\tau = [(i, r), j]$ equals the minimum of the number of RBC units supplied by vertex $(i, r) \in \mathcal{V}_1$ and the number of RBC units demanded by vertex $j \in \mathcal{V}_2$. Hence, the capacity of edge $\tau \in \mathcal{T}$ is equal to $u_\tau = \min\{s_{ir}, d_j\}$. Finally, the cost for allocating one RBC unit to an edge $\tau = [(i, r), j]$ is given by c_τ.

Let $\boldsymbol{x} \in \mathcal{X}$ be an allocation between the RBC units in inventory and the RBC units requested, where

$$
\mathcal{X} = \left\{ \boldsymbol{x} \in \mathbb{N}_0^{|\mathcal{T}|} \ \middle| \ \begin{array}{ll} \displaystyle\sum_{j \in \mathcal{V}_2} x_{[(i,r),j]} \leq s_{ir}, & \forall\, (i, r) \in \mathcal{V}_1, \ \text{(supply constraints)} \\[2mm] \displaystyle\sum_{(i,r) \in \mathcal{V}_1} x_{[(i,r),j]} \leq d_j, \forall\, j \in \mathcal{V}_2, & \text{(demand constraints)} \\[2mm] 0 \leq x_\tau \leq u_\tau, & \forall\, \tau \in \mathcal{T}, \quad \text{(capacity constraints)} \end{array} \right\}
$$
(6.2)

represents the set of feasible allocations (i.e., an allocation $\boldsymbol{x}$ should satisfy the supply, demand, and capacity constraints). The goal is to find an allocation $\boldsymbol{x} \in \mathcal{X}$ that maximizes the number of outstanding requests that can be satisfied with compatible RBC units, but has minimum costs among those maximum flow allocations:

$$
\min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \sum_{\tau \in \mathcal{T}} c_\tau x_\tau \ \middle| \ \sum_{\tau \in \mathcal{T}} x_\tau = \max_{\boldsymbol{y} \in \mathcal{X}} \left\{ \sum_{\tau \in \mathcal{T}} y_\tau \right\} \right\}.
$$
(6.3)
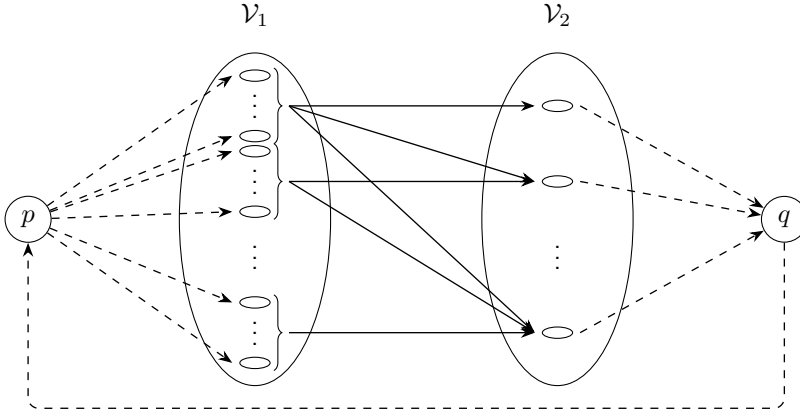
Figure 6.4: Graph $\mathcal{G}_{pq}(\mathcal{V}_{pq}, \mathcal{T}_{pq})$ underlying the circulation problem, with vertex set $\mathcal{V}_{pq} = \{p\} \cup \mathcal{V} \cup \{q\}$ and edge set $\mathcal{T}_{pq} = \mathcal{T}_p \cup \mathcal{T} \cup \mathcal{T}_q \cup \{[q, p]\}$. $\mathcal{V}_1$ represents the inventory of RBC units and $\mathcal{V}_2$ the blood groups requested.

This optimization problem is known as a minimum cost flow problem (MCFP), which can be solved efficiently due to its specific structure [9, 128].

Solving the minimum cost flow problem

To find an optimal allocation between the RBC units in inventory and the RBC units requested, or more mathematically, a maximum flow from the set of sources $\mathcal{V}_1$ to the set of sinks (destinations) $\mathcal{V}_2$ with minimum costs, the inventory allocation problem is transformed into a circulation problem (Figure 6.4). First, the vertex set of $\mathcal{G}$ is extended with a source vertex p and a sink vertex q. Second, two new edge sets are introduced: $\mathcal{T}_p$ and $\mathcal{T}_q$. $\mathcal{T}_p = \{\tau \mid \tau = [p, (i, r)], (i, r) \in \mathcal{V}_1\}$ consists of edges from p to all vertices in $\mathcal{V}_1$. The edges $\tau \in \mathcal{T}_p$ have capacity $u_\tau = s_{ir}$ and unit cost $c_\tau = 0$. $\mathcal{T}_q = \{\tau \mid \tau = [j, q], j \in \mathcal{V}_2\}$ consists of edges from all vertices in $\mathcal{V}_2$ to q. The edges $\tau \in \mathcal{T}_q$ have capacity $u_\tau = d_j$ and unit cost $c_\tau = 0$. Third, an edge $[q, p]$ from the sink node to the source node is constructed. This edge might have infinite capacity, but for computational convenience the capacity of this edge can be restricted by the number of units requested, i.e., $u_{[q,p]} = \sum_{j \in \mathcal{V}_2} d_j$. The unit cost of edge $[q, p]$ is set equal to $c_{[q,p]} = -\max_{\tau \in \mathcal{T}} \{c_\tau\} - 1$. Summarizing, the graph $\mathcal{G}_{pq}(\mathcal{V}_{pq}, \mathcal{T}_{pq})$ underlying the circulation problem has vertex set $\mathcal{V}_{pq} = \{p\} \cup \mathcal{V} \cup \{q\}$ and edge set $\mathcal{T}_{pq} = \mathcal{T}_p \cup \mathcal{T} \cup \mathcal{T}_q \cup \{[q, p]\}$, where each edge $\tau \in \mathcal{T}_{pq}$ has capacity u_τ and cost c_τ.

By transforming the inventory allocation problem into a circulation problem (a network flow problem without any source or sink vertices, but only transshipment vertices), a new set of constraints has to be introduced. These constraints ensure that the amount of flow that enters a vertex $w \in \mathcal{V}_{pq}$ is equal to the amount of flow that leaves this vertex. As the capacity constraints remain, the set of feasible

solutions to the circulation problem is given by

$$
\mathcal{X}_{pq} = \left\{ \boldsymbol{x} \in \mathbb{N}^{|\mathcal{T}_{pq}|} \;\middle|\; \sum_{\tau \in \mathcal{T}_{pq}^{\text{in}}(w)} x_\tau = \sum_{\tau \in \mathcal{T}_{pq}^{\text{out}}(w)} x_\tau, \quad \forall\, w \in \mathcal{V}_{pq} \\ 0 \leq x_\tau \leq u_\tau, \qquad\qquad \forall\, \tau \in \mathcal{T}_{pq} \right\}
\tag{6.4}
$$

where $\mathcal{T}_{pq}^{\text{in}}(w) = \{\tau \in \mathcal{T}_{pq} \mid \tau = [v, w],\ v \in \mathcal{V}_{pq}\}$ represents the set of edges that enter a vertex $w \in \mathcal{V}_{pq}$ and $\mathcal{T}_{pq}^{\text{out}}(w) = \{\tau \in \mathcal{T}_{pq} \mid \tau = [w, v],\ v \in \mathcal{V}_{pq}\}$ represents the set of edges that leave this vertex. A minimum cost circulation can be found by solving the following optimization problem:

$$
\min_{\boldsymbol{x} \in \mathcal{X}_{pq}} \left\{ \sum_{\tau \in \mathcal{T}_{pq}} c_\tau x_\tau \right\}.
\tag{6.5}
$$

Remark 6.1 (LP-relaxation). As the capacities u_τ are integer for all $\tau \in \mathcal{T}_{pq}$, the LP-relaxation of optimization problem (6.5) has an integer optimal solution [9, 128]. Therefore, the integrality constraints can be omitted. This substantially enhances the practical applicability, as computational experiments have shown that an optimal solution to the LP-relaxation of the inventory allocation problem can be found within seconds.

6.3.3 Issuing policy

The rareness of a blood group is based on prevalence of a blood group in the general population, often measured as a percentage (0.1%), or as a numerical rate (1 in 1000). However, the likelihood that an RBC unit can be used issued before it expires, also depends on the prevalence of compatible blood groups in the transfusion recipient population. Similarly, the likelihood that a request for an RBC unit can be satisfied, also depends on the prevalence of compatible blood groups in the donor population. Therefore, two new measures are introduced: the usability factor and the receivability factor.

Definition 6.1 (Usability factor). The usability factor of an RBC unit with blood group $i \in \mathcal{B}$ is defined as

$$
F_i^s = \sum_{j:i \leq j} f_j,
\tag{6.6}
$$

with f_j^d the proportion of RBC requested with blood group $j \in \mathcal{B}$.
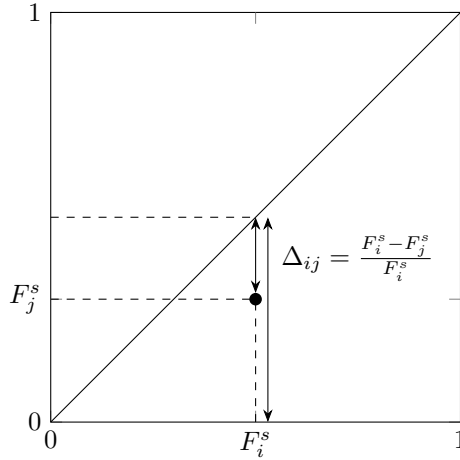
Figure 6.5: Relative opportunity loss (Δ_{ij}) of a match $\tau = [(i,r),j]$, with F_i^s the usability factor the RBC unit supplied and F_j^s the usability factor of the RBC unit requested.

Definition 6.2 (Receivability factor). The receivability factor of an RBC unit with blood group $j \in \mathcal{B}$ is defined as

$$F_j^d = \sum_{i:i \leq j} f_i, \tag{6.7}$$

with f_i^s the proportion of RBC supplied with blood group $i \in \mathcal{B}$.

The proposed issuing policy has to find a trade-off between 'relative opportunity loss', as specified below, and the age of the issued RBC unit. Therefore let us first define the concept of 'relative opportunity loss'.

Definition 6.3 (Relative opportunity loss). Let F_i^s be the usability factor of the RBC unit supplied and let F_j^s be the usability factor of the RBC unit requested. The relative opportunity loss of a match $\tau = [(i,r),j]$ is given by

$$\Delta_{ij} = \frac{F_i^s - F_j^s}{F_i^s}. \tag{6.8}$$

It indicates the potential loss of an RBC unit with blood group i is issued to satisfy a request for an RBC unit with blood group j (Figure 6.5).

In the illustrative example from Section 6.3.1 we used a linear cost function for the allocation of an RBC unit to an edge $\tau = [(i,r),j]$ of the form $c_\tau = \gamma \cdot \Delta_{ij} + (1-\gamma) \cdot \frac{r}{R}$. This would imply a linear trade-off between the relative opportunity loss and age. However, the relation between costs and relative opportunity loss or age might have any form, with function $g_1(\Delta)$ and $g_2(r)$

$$c_\tau(\Delta_{ij}, r) = \gamma g_1(\Delta_{ij}) + (1-\gamma) g_2(r) \tag{6.9}$$

where $g_1(\Delta_{ij})$ and $g_2(r)$ should be seen as functions to reflect the quality of the match and the outdating respectively. Note that the choice of functions g_1 and g_2 does not influence the solvability of the inventory allocation problem as the objective function remains linear in x. To be able to let the effect of either the matching quality or future consequences be more distinctive (or dominant), in the numerical experiments that will be discussed later we choose the form

$$g_1(\Delta_{ij}) = 1 - \exp(-\alpha \cdot \Delta_{ij}) \text{ and } g_2(r) = 1 - \exp(-\beta \cdot \frac{r}{R}). \qquad (6.10)$$

where the parameters α, β can be interpreted as discount factors. For the computational experiments we will take $\alpha = \beta = 7\log(2)$. This would roughly imply that the costs double for any identical RBC unit that is one week older.

Remark 6.2 (Cost function). Clearly, the specific choice of cost functions is arbitrary. Nevertheless, the results in Section 6.5.2 as based on the MCFP optimization with these cost functions show that this function performs well in practice.

6.4   Evaluation model

In the previous section, the inventory allocation problem has been described as an MCFP, which finds an optimal allocation between RBC units in inventory and the RBC units requested. This allocation is optimal when it minimizes the objective function, which represents a static issuing policy as discussed in Section 6.3.3. However, so far no stochastic effects with respect to the supply and demand of RBC units have been taken into account The incorporation of these effects will be discussed in this section.

6.4.1   Simulation setup

A simulation study is conducted in which the optimal deterministic MCFP solutions are incorporated in a stochastic setting, i.e., the deterministic inventory allocation model from Section 6.3.2 is integrated with a stochastic supply and demand of RBC units. Its performance is evaluated over time. A schematic representation of the setup of the simulation is given in Figure 6.6. The stochasticity incorporated involves:

1. requests for particular blood groups,

2. response rates of donors,

3. lead time distributions,

4. distribution of antigens within the donor population.
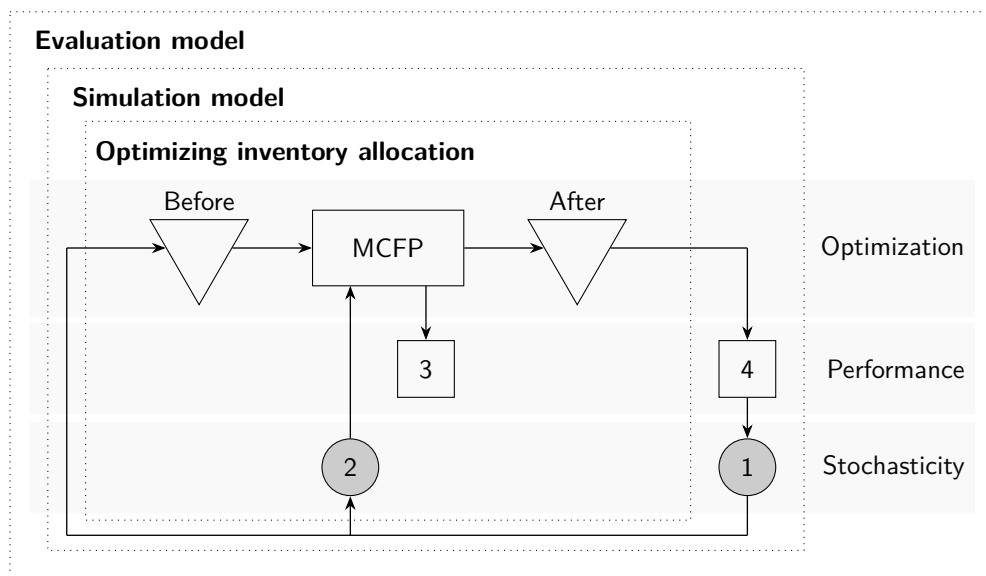
Figure 6.6: Schematic representation of the simulation model. It incorporates the stochasticity in the supply (1) and demand (2) of RBC units, and evaluates the impact of the issuing policy on the percentage of units short (3) and the percentage of units outdated (4).

Sections 6.4.2 and 6.4.3 discuss how historical data on supply and demand of RBC units was obtained from the central information system of the Dutch blood bank (*eProgesa*) for each of these random features.

6.4.2 Demand distributions

A database containing information on all requested RBC units between 01/01/2014 and 31/12/2016 was extracted from the Dutch blood bank information system *eProgesa*. A single record (or row) in this database indicates the amount of RBC units requested with a particular blood group on a specific date by a hospital. Records for which no ABO, D blood group was specified (1.3%), or had a date that coincided with a national holiday (1.3%) were excluded. The remaining records (97.3%) were divided into 56 groups based on the eight ABO, D blood groups, and seven days of the week (i.e., Monday, Tuesday, ..., Sunday). Within each group, all RBC units that were requested on the same date were aggregated, resulting in 56 empirical demand distributions consisting of approximately 156 data points (i.e., number of Mondays, Tuesdays, etc. in the period 01/01/2014 to 31/12/2016).

Each empirical demand distribution was approximated by a discrete distribution based on the fitting procedure of Adan et al. [2], which preserves the mean and coefficient of variation of the empirical distribution. This implied that each

empirical distribution could be approximated by a mixture of two negative binomial distributions (Supplemental Figure 6.8). Based upon the prevalence of non-ABO, D antigens in the Caucasian population [94] an extended blood group was assigned.

6.4.3   Supply distributions

A databases containing information on all invitations and donations between 01/01/2014 and 31/12/2016 were extracted from the Dutch blood bank information system *eProgesa*. Each donations was linked to an invitation based on donor identification number. Donations during weekend days (0.03%) and with a lead time of more than four weeks (3.7%) were excluded. The average response rate was equal to 0.4 and independent of the ABO, D blood group of the donor. Lead time distributions were computed by counting the number of days between donation and invitation.

6.5   Computational experiments and results

This section describes the experiments and results that were conducted by the integrated approach presented in Section 6.4 to analyze the long-term performance of the proposed issuing policy, which takes both the age and rareness of the RBC units in inventory into account. Section 6.5.1 first discusses the setup of the experiments followed by the computational results in Section 6.5.2.

6.5.1   Simulations performed

Each experiment starts with the initialization of the simulation model. A run spans over one year (365 days). This is preceded by 56 days establish an equilibrium in the model states. Starting values form a uniformly distributed inventory from 0 days to k weeks old, where k is the shelf life parameter. For the first and second antigen sets, 100 runs are used to compute the average performance and the confidence intervals, for the third antigen set, only 1 run was used to reduce computation time. This results from the fact that thousands of decisions have to be taken. The decision at one decision moment, as would be required in practice, still only requires a few seconds. We have analyzed three different antigen sets:

- $\mathcal{A}_1 = \{\mathsf{A}, \mathsf{B}, \mathsf{D}\}$,

- $\mathcal{A}_2 = \{\mathsf{A}, \mathsf{B}, \mathsf{D}, \mathsf{C}, \mathsf{c}, \mathsf{E}, \mathsf{e}, \mathsf{K}\}$,

- $\mathcal{A}_3 = \left\{\mathsf{A}, \mathsf{B}, \mathsf{D}, \mathsf{C}, \mathsf{c}, \mathsf{E}, \mathsf{e}, \mathsf{K}, \mathsf{Fy}^{\mathsf{a}}, \mathsf{Fy}^{\mathsf{b}}, \mathsf{Jk}^{\mathsf{a}}, \mathsf{Jk}^{\mathsf{b}}, \mathsf{S}, \mathsf{s}\right\}$

The other parameters that are varied in the evaluation model are the maximum residual shelf life of an RBC unit ($R = 14, 21, 28$ or 35 days), the size of the

inventory ($k = 0.5, 1, 1.5$ and 2 weeks), and the parameter indicating the trade off between the rareness and age of the RBC unit supplied ($\gamma = 0, 1,$ and $\frac{1}{2}$).

The inventory allocation model from Section 6.3 and the simulation model from Section 6.4 have both been implemented in Matlab R2017b and solved on a laptop with an Intel Core i5-4210U CPU and 8GB of RAM. In each iteration of the simulation model, an optimal allocation between the RBC units in inventory and the RBC units requested was determined using Gurobi 8.1.0 (a tool for solving linear programming problems). As this optimal allocation was found within seconds for all investigated parameter settings, the inventory allocation model might be converted into a practical decision making tool to support real-time matching.

6.5.2　Computational results and analyses

The number of units short and the number of units outdated are described for all three antigen sets, all combinations of the inventory parameters R (maximum residual shelf life), k (inventory size in weeks), and the cost parameter γ. Sections 6.5.2 and 6.5.2 are concentrated on three specific scenarios ($R = 35$, $k = 1$, $\gamma = 0, 1, \frac{1}{2}$), which is most similar to the current inventory management. In the following sections we will investigate the impact of γ on the percentage of requests that can be satisfied with antigen identical RBC units and the average age of RBC units when issued.

Shortage versus outdating

Both shortage (η) and outdating (ω) are very small in all scenarios. There is no shortage for any of the runs with antigen sets $\mathcal{A}_1$. The results for antigen sets $\mathcal{A}_2$ and $\mathcal{A}_3$ are given in Table 6.2 and 6.3. For the runs with antigen set $\mathcal{A}_2$ there is some shortage, but this is at most 0.1%. Interestingly, even when an inventory of only $k = 0.5$ weeks is kept. Outdating is dependent on k, and becomes problematic when the maximum residual shelf life becomes $R = 14$ days and more than one week of inventory is kept. Outdating is negligible for a maximum residual shelf of 21, 28, and 35 days.

For antigen set $\mathcal{A}_3$ the results become more interesting and it seems that there is a choice to be made: when the inventory size k is 0.5 weeks, shortages become more than 2% for almost all parameter settings. However, shortage decrease to less than 1% if 2 weeks of inventory is kept. The drawback of two weeks of inventory is that it increases the level of outdating, which is problematic if the maximal residual shelf life (R) is small. However for $R = 35$ a relatively low outdating rate is found, if a FIFO issuing policy is applied (in the order of 0.5% up to 1.5%).

Looking at the impact of parameter γ that balances the preference between age and rareness of blood product, for $\gamma = 0$ and $\gamma = 0.5$ we see that the issuing policy that combines both age and rareness of blood product performs only slightly

Table 6.2: Percentage of unit short (η) and outdated (ω) for antigen set $\mathcal{A}_2 = \{A, B, D, C, c, E, e, K\}$.

R	k	FIFO ($\gamma = 0$) η	FIFO ($\gamma = 0$) ω	MROL ($\gamma = 1$) η	MROL ($\gamma = 1$) ω	FIFO / MROL ($\gamma = \frac{1}{2}$) η	FIFO / MROL ($\gamma = \frac{1}{2}$) ω
14 days	0.5 week	0.1%	0.2%	0.1%	0.4%	0.1%	0.2%
	1.0 week	0.0%	0.7%	0.0%	1.0%	0.0%	0.8%
	1.5 week	0.0%	4.4%	0.0%	4.8%	0.0%	4.7%
	2.0 week	0.0%	11.9%	0.0%	12.4%	0.0%	12.4%
21 days	0.5 week	0.1%	0.1%	0.1%	0.2%	0.1%	0.1%
	1.0 week	0.0%	0.2%	0.0%	0.3%	0.0%	0.2%
	1.5 week	0.0%	0.3%	0.0%	0.5%	0.0%	0.4%
	2.0 week	0.0%	1.3%	0.0%	1.7%	0.0%	1.5%
28 days	0.5 week	0.1%	0.0%	0.1%	0.1%	0.1%	0.1%
	1.0 week	0.0%	0.1%	0.0%	0.2%	0.0%	0.1%
	1.5 week	0.0%	0.1%	0.0%	0.2%	0.0%	0.1%
	2.0 week	0.0%	0.2%	0.0%	0.3%	0.0%	0.2%
35 days	0.5 week	0.1%	0.0%	0.1%	0.1%	0.1%	0.0%
	1.0 week	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
	1.5 week	0.0%	0.1%	0.0%	0.1%	0.0%	0.1%
	2.0 week	0.0%	0.1%	0.0%	0.2%	0.0%	0.1%

worse with respect to outdating for all scenarios investigated. Similarly, comparing the values for the parameters $\gamma = 1$ and $\gamma = 0.5$, we see that the combined issuing policy also performs only slightly worse with respect to shortage. Hence, the currently applied cost function that takes both the age and rareness of a RBC units into account seems to perform very well.

Percentage of requests satisfied with antigen identical red blood cell unit

Let us now restrict ourselves to only 3 out of the 144 different scenarios by fixing the set of antigens considered, the maximum residual shelf life, and the inventory size and only differ in the cost parameter γ. Let us divide the different blood groups into five categories based on their receivability factor (Definition 2). Table 6.4 shows the results for the percentage of request that could be satisfied with an antigen identical RBC unit when looking at the receivability factor of the requested blood groups. From this table, it can be concluded that if the receivability factor decreases (RBC units become more rare), it is less likely that identical RBC units can be issued. Moreover, the probability that a request cannot be satisfied with either an identical or a compatible blood group increases with the rareness of the RBC unit. The proportion of requests that can be satisfied with an antigen identical

Table 6.3: Percentage of unit short (η) and outdated (ω) for antigen set $\mathcal{A}_3 = \left\{\text{A}, \text{B}, \text{D}, \text{C}, \text{c}, \text{E}, \text{e}, \text{K}, \text{Fy}^a, \text{Fy}^b, \text{Jk}^a, \text{Jk}^b, \text{S}, \text{s}\right\}$.

R	k	FIFO ($\gamma = 0$) η	ω	MROL ($\gamma = 1$) η	ω	FIFO / MROL ($\gamma = \frac{1}{2}$) η	ω
14 days	0.5 week	2.9%	2.3%	1.8%	5.8%	2.1%	2.8%
	1.0 week	1.4%	4.6%	0.8%	9.8%	0.9%	5.4%
	1.5 week	0.7%	9.4%	0.4%	14.3%	0.4%	10.1%
	2.0 week	0.3%	15.4%	0.2%	19.1%	0.2%	16.3%
21 days	0.5 week	3.3%	1.2%	2.4%	3.2%	2.4%	1.2%
	1.0 week	1.7%	1.9%	1.1%	5.7%	1.3%	2.1%
	1.5 week	1.1%	3.2%	0.6%	7.3%	0.7%	3.5%
	2.0 week	0.6%	5.6%	0.4%	10.5%	0.4%	6.2%
28 days	0.5 week	3.6%	0.8%	2.4%	2.1%	2.8%	0.7%
	1.0 week	2.1%	1.1%	1.3%	3.3%	1.5%	1.2%
	1.5 week	1.1%	1.5%	0.8%	4.7%	0.9%	1.6%
	2.0 week	0.9%	2.5%	0.4%	5.9%	0.5%	2.6%
35 days	0.5 week	3.7%	0.5%	2.7%	1.5%	2.8%	0.5%
	1.0 week	2.2%	0.7%	1.2%	2.2%	1.3%	0.8%
	1.5 week	1.4%	0.9%	0.9%	3.2%	0.8%	1.0%
	2.0 week	1.0%	1.4%	0.5%	3.7%	0.5%	1.5%

RBC unit increases if more weight is put on minimizing the relative opportunity loss, which is exactly what we would expect. As already said in the previous section, also the percentage of RBC units short decrease when more weight is put on minimizing the relative opprtunity loss. However, the difference between $\gamma = \frac{1}{2}$ and $\gamma = 1$ is small (0.1%).

Table 6.5 shows the results for the percentage of requests that could be satisfied with an antigen identical RBC unit, but now looking at the usability factor of the issued RBC unit (Definition 1). In general it holds that rare RBC units have a high usability factor, as they are negative for many antigens and therefore compatible with many blood groups. Surprisingly, the percentage of identical matches decreases when RBC units are less rare. The percentage of units outdated increases when the usability factor decreases, which is exactly what we would expect. The FIFO issuing policy has the lowest percentage of units outdated, but the difference with the combined issuing policy is small. When comparing a FIFO policy ($\gamma = 1$) with the combined FIFO/MROL policy ($\gamma = \frac{1}{2}$) we see that the percentage of requests that can be satisfied with identical RBC units is higher for the combined issuing policy. This is due to the fact that 1.4% of outdating is prevented by issuing older units earlier.

Table 6.4: All RBC units requested were classified into five based on their receivability factor. For each requested RBC unit we investigated whether the issued RBC unit was identical, compatible or not available.

(a) FIFO ($\gamma = 0$)

receivability	identical		compatible		shortage		
$[10^{-0}, 10^{-1})$	14.2%	(88.2%)	1.9%	(11.8%)	0.0%	(0.0%)	16.1%
$[10^{-1}, 10^{-2})$	50.6%	(89.8%)	5.6%	(10.0%)	0.1%	(0.2%)	56.4%
$[10^{-2}, 10^{-3})$	21.4%	(81.5%)	3.3%	(12.4%)	1.6%	(6.1%)	26.3%
$[10^{-3}, 10^{-4})$	0.6%	(49.4%)	0.2%	(12.0%)	0.5%	(38.6%)	1.3%
$[10^{-4}, 10^{-5})$	0.0%	(11.9%)	0.0%	(0.0%)	0.0%	(88.1%)	0.0%
	86.8%		10.9%		2.2%		100%

(b) MROL ($\gamma = 1$)

receivability	identical		compatible		shortage		
$[10^{-0}, 10^{-1})$	15.7%	(97.7%)	0.4%	(2.3%)	0.0%	(0.0%)	16.1%
$[10^{-1}, 10^{-2})$	54.2%	(96.1%)	2.2%	(3.8%)	0.0%	(0.0%)	56.4%
$[10^{-2}, 10^{-3})$	23.2%	(88.3%)	2.3%	(8.6%)	0.8%	(3.0%)	26.3%
$[10^{-3}, 10^{-4})$	0.7%	(56.1%)	0.2%	(12.8%)	0.4%	(31.1%)	1.3%
$[10^{-4}, 10^{-5})$	0.0%	(10.4%)	0.0%	(9.0%)	0.0%	(80.6%)	0.0%
	93.8%		5.0%		1.2%		100%

(c) FIFO / MROL ($\gamma = \frac{1}{2}$)

receivability	identical		compatible		shortage		
$[10^{-0}, 10^{-1})$	15.4%	(95.8%)	0.7%	(4.2%)	0.0%	(0.0%)	16.1%
$[10^{-1}, 10^{-2})$	54.0%	(95.8%)	2.4%	(4.2%)	0.0%	(0.0%)	56.4%
$[10^{-2}, 10^{-3})$	23.3%	(88.6%)	2.1%	(8.0%)	0.9%	(3.4%)	26.3%
$[10^{-3}, 10^{-4})$	0.7%	(57.1%)	0.1%	(11.5%)	0.4%	(31.5%)	1.3%
$[10^{-4}, 10^{-5})$	0.0%	(10.4%)	0.0%	(0.0%)	0.0%	(89.6%)	0.0%
	93.4%		5.3%		1.3%		100%

Issuing age

Figure 6.7 shows box plots for the issuing age of RBC units stratified by both their usability and receivability factor. Note that on the left y-axis the age the issued RBC unit is displayed and on the right y-axis the residual shelf life ($= 35 -$ age). Looking at Figure 6.7a we see that RBC units with a high usability factor are issued at a younger age than RBC units with a low usability factor. RBC units with a

Table 6.5: All RBC units in inventory were classified into five based on their usability factor. For each RBC unit in inventory we investigated whether the it was eiter issued identical, compatible or expired.

(a) FIFO ($\gamma = 0$)

usability	identical		compatible		outdated		
$[10^{-0}, 10^{-1})$	29.0%	(94.7%)	1.6%	(5.3%)	0.0%	(0.0%)	30.6%
$[10^{-1}, 10^{-2})$	50.7%	(89.9%)	5.8%	(10.1%)	0.0%	(0.0%)	56.4%
$[10^{-2}, 10^{-3})$	8.2%	(68.8%)	3.3%	(27.6%)	0.4%	(3.6%)	11.9%
$[10^{-3}, 10^{-4})$	0.3%	(30.0%)	0.5%	(46.5%)	0.2%	(23.4%)	1.0%
$[10^{-4}, 10^{-5})$	0.0%	(0.0%)	0.0%	(53.6%)	0.0%	(46.4%)	0.0%
	88.2%		11.1%		0.7%		100%

(b) MROL ($\gamma = 1$)

usability	identical		compatible		outdated		
$[10^{-0}, 10^{-1})$	30.0%	(97.2%)	0.6%	(2.1%)	0.2%	(0.7%)	30.8%
$[10^{-1}, 10^{-2})$	52.9%	(94.2%)	2.3%	(4.1%)	1.0%	(1.7%)	56.2%
$[10^{-2}, 10^{-3})$	9.5%	(78.6%)	1.7%	(14.2%)	0.9%	(7.2%)	12.1%
$[10^{-3}, 10^{-4})$	0.2%	(53.7%)	0.3%	(28.0%)	0.2%	(18.2%)	0.9%
$[10^{-4}, 10^{-5})$	0.0%	(15.0%)	0.0%	(46.7%)	0.0%	(38.3%)	0.0%
	92.9%		4.9%		2.2%		100%

(c) FIFO / MROL ($\gamma = \frac{1}{2}$)

usability	identical		compatible		outdated		
$[10^{-0}, 10^{-1})$	30.3%	(97.6%)	0.7%	(2.4%)	0.0%	(0.0%)	31.0%
$[10^{-1}, 10^{-2})$	53.5%	(95.2%)	2.6%	(4.7%)	0.1%	(0.1%)	56.2%
$[10^{-2}, 10^{-3})$	9.7%	(81.2%)	1.7%	(14.5%)	0.5%	(4.3%)	11.9%
$[10^{-3}, 10^{-4})$	0.5%	(57.1%)	0.2%	(26.0%)	0.2%	(16.8%)	0.9%
$[10^{-4}, 10^{-5})$	0.0%	(13.0%)	0.0%	(40.6%)	0.0%	(46.4%)	0.0%
	93.9%		5.3%		0.8%		100%

usability factor between 0.1 and 1 are issued within one week, when a FIFO issuing policy is used and within two weeks when an MROL issuing policy is used. This increase can be explained by the fact that with an MROL issuing policy rare units will remain longer in stock.

In Figure 6.7b blood groups were classified on the receivability factor of the units requested. From the figure it is clear that the age of the RBC units issued decreases with a decreasing receivability factor of the requested RBC units.

Figure 6.7: Issuing age of the RBC units in inventory, classified on blood group identical and blood group compatible matches.

(a) Blood groups $i \in \mathcal{B}$ were aggregated based on usability factor (F_i^s) of the RBC units issued.
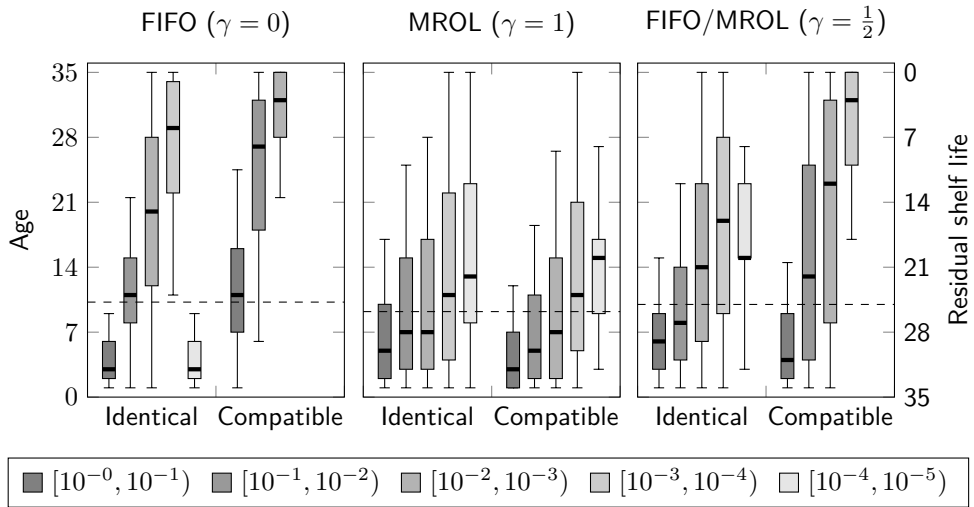
(b) Blood groups $j \in \mathcal{B}$ were aggregated based on receivability factor (F_j^d) of the RBC units requested.
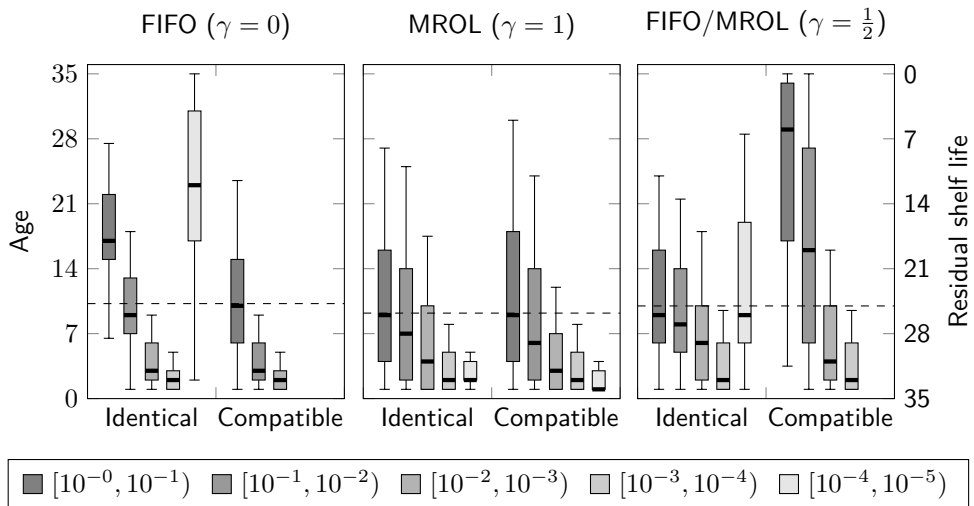
6.6   Conclusions

This chapter presents a mathematical framework that can be used to evaluate the long-term performance of a predetermined issuing policy, which prioritizes between the age and rareness of the RBC units in inventory. Daily decisions about which RBC units to select from inventory to satisfy requests from hospitals are optimized by modeling the inventory allocation problem as an MCFP. A simulation study, which incorporates the stochasticity in the supply and demand of RBC units, is conducted to assess the impact of these (optimal) daily decisions on the percentage of units short and percentage of units outdated.

Currently the selection of RBC units to be issued is a manual process. This will no longer be possible when the number of antigens that have to matched is extended, as the inventory will rapidly become far too complex for a human to manage. Hence, a system that is able to provide a fast and smart selection of RBC units to be issued is urgently needed. The model presented in this chapter seems appropriate to fulfill this need. More specifically, the inventory allocation model, presented in Section 6.3, can be seen as the engine of a decision support system that guides the issuance of extensively matched RBC units.

The proposed issuing policy that is investigated in this chapter, based on a specific choice of the cost function, seems to perform well with respect to the percentage of units short and percentage of units outdated. However, it does not explicitly take potential future shortage or outdating into account. By a modelling this problem as a Markov Decision Process, as was done in Chapter 5, the decision concerning which RBC units to select from inventory can be optimized exactly. The drawback of this approach is that the number of states (inventory composition) grows exponentially with the number of RBC units in inventory considered, which prevents the application of this approach in practice. Our results show that the mathematical framework proposed in this chapter performs well with respect to both the percentage of units short and percentage of units outdated.
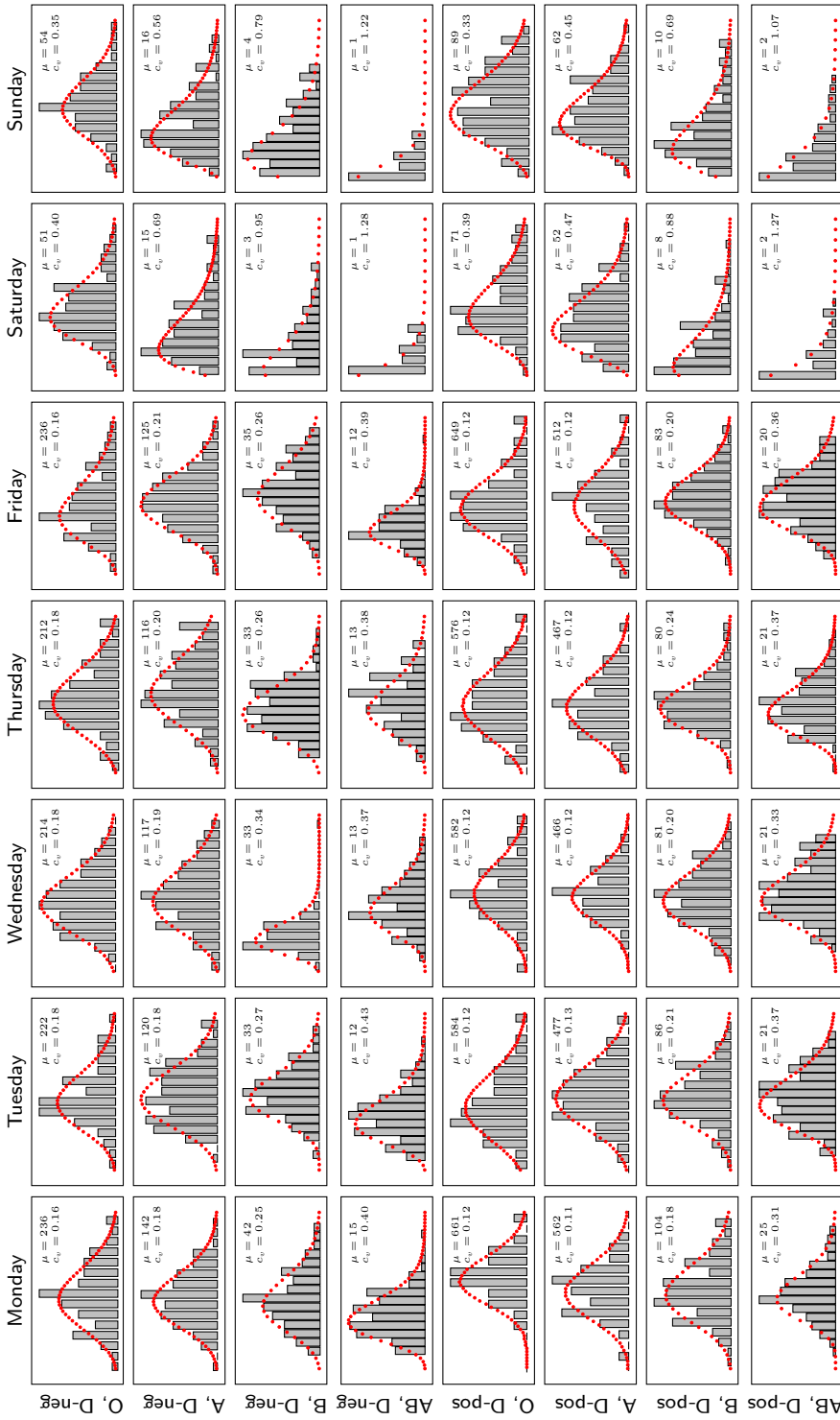
Figure 6.8: Empirical demand distributions for the number of RBC units requested with a particular ABO, D blood group on a specific day of the week. Each empirical demand distribution was approximated by a mixture of two negative binomial distribution, such that the mean (μ) and coefficient of variation (c_v) of the original empirical distribution were preserved.

Mathematical optimization for alloimmunization prevention

7.1 Introduction

The most common adverse event associated with blood transfusion is the formation of alloantibodies due to incompatibilities between the red blood cell (RBC) antigens of donors and transfusion recipients. Alloantibodies may cause acute or delayed hemolytic transfusion reactions during subsequent transfusions and may lead to hemolytic disease of the foetus and newborn (HDFN) [27, 58, 96, 113]. Alloimmunization can be prevented by matching RBC antigens of donors and transfusion recipients, however this requires that antigen profiles are known in both populations and compatible RBC units are available [38, 65].

At present, antigen typing is mainly done by serology and only a limited number of antigens can be tested. In contrast, genotyping can correctly predict the expression of hundreds of antigens in a single assay. In the last decade, several high throughput platforms have been developed and molecular profiling of donors and transfusion recipients is likely to become common practice [20, 21, 40, 57, 74, 75, 122, 126].

Currently, extended preventive matching strategies are only applied for specific groups of transfusion recipients and include a limited number of antigens. In the Netherlands, women under 45 years of age receive c, E, and K compatible units; individuals with myelodysplastic syndromes, alloantibodies, and autoantibodies, receive at least C, c, E, e, and K compatible units; and individuals with sickle cell disease and thalassemia are in addition matched for the antigens Fy[a], (and if possible for) Jk[b], S, and s [18]. Preventive matching strategies have led to a notable reduction in the formation of alloantibodies [19, 70, 81, 104].

The availability of completely matched RBC units for all transfusion recipients would eliminate transfusion-induced alloimmunization. However, given the large number of antigens, this is thought to be infeasible in practice, even when consid-

ering only the most clinically relevant antigens. RBC inventory sizes in blood banks and hospitals are finite, which limits the likelihood that matched RBC units are available. The feasibility of preventive matching when RBC antigens of donors and transfusion recipients would be comprehensively typed has never been determined.

The goal of this study is to compute the proportion of alloimmunization that can be prevented, when all donors and recipients are typed for A, B, D, and twelve other minor blood group antigens, together responsible for the far majority of alloimmunization events [36]. This includes an optimal sequence for antigen matching, when the inventory contains an insufficient number of RBC units to match a transfusion recipient on all fifteen antigens considered.

The current assumption in transfusion medicine is that there are numerous logistical and financial challenges when complete donor-recipient antigen matching is attempted. Therefore, the next best alternative is to apply preventive matching only for those transfusion recipients that are at high risk for alloimmunizaiton. By mathematical modeling, we will show that preventive matching is feasible for all transfusion recipients and not only for those at high risk. Our study paves the way for a shift in transfusion policy from preventing the consequences of alloimmunization by screening for alloantibodies to preventing alloimmunization by matching.

7.2   Study design and methods

To quantify and optimize the proportion of alloimmunization that can be prevented by extending matching strategies up to fifteen antigens, information is required concerning

1) the probability that a recipient would form antibodies against particular antigens

2) the availability of extensively matched RBC units.

7.2.1   Antibody formation

Estimates on RBC alloimmunization risks against the selected antigens (C, c, E, e, K, Fy^a, Fy^b, Jk^a, Jk^b, M, S, and s) in an incident new-user cohort of previously non-transfused, non-alloimmunized Caucasian transfusion recipients receiving non-extended matched RBC transfusions are used to calculate the proportion of transfusion-induced alloantibodies formed against these antigens (Table 7.1) [36]. These proportions are calculated by dividing the number of alloimmunizations against a selected antigen by the total number of alloimmunizations against

Table 7.1: Specificity and proportion (p_a) of selected first-time formed clinically relevant alloantibodies.

Alloantibody	anti-E	anti-K	anti-Jk[a]	anti-c	anti-Fy[a]	anti-C	anti-M
p_a	0.373	0.257	0.105	0.078	0.051	0.046	0.038
Alloantibody	anti-S	anti-Jk[b]	anti-Fy[b]	anti-e	anti-s		total
p_a	0.017	0.015	0.011	0.008	0		1

all selected antigens.

7.2.2   Availability of matched RBC units

Three inventory allocation models with increasing levels of complexity are applied to quantify the likelihood that RBC units with a particular blood group are present in inventory. The models differ in inventory composition (i.e., dynamics between the distribution of blood groups in the donor and transfusion recipient population) and issuing policy (Table 7.2). The distribution of blood groups in the donor and transfusion recipient populations are assumed to be equal to the distribution of blood group in a Caucasian population. They are obtained by combining the antigen frequency tables of the ABO, Rhesus, Kell, Duffy, Kidd, and MNS blood group systems (as presented in Section 4.4).

Inventory allocation model 1 is based on two assumptions:

1) the composition of the inventory is a reflection of the distribution of blood groups in the donor population and remains unchanged after issuing RBC units (static inventory composition),

2) any compatible RBC unit can be issued (antigen compatible issuing policy).

The second assumption implies that a transfusion recipient, who is positive for particular minor blood group antigens might be transfused with blood from a donor who is negative for these antigens. As a result, RBC units that are negative for many antigens will be issued more frequently than can be expected from their occurrence in the donor population. Inventory allocation model 1, therefore, overestimates the practical matching capacity and will not accurately reflect the likelihood of having compatible RBC units in inventory.

Inventory allocation model 2 differs from model 1 in that it applies an antigen identical issuing policy instead of an antigen compatible issuing policy, i.e., RBC

Table 7.2: Inventory allocation models used to quantify the likelihood that RBC units are available from inventory. The inventory allocation models differ with respect to inventory composition and issuing policy.

	Inventory composition	Issuing policy
Inventory allocation model 1	Static	Antigen compatible
Inventory allocation model 2	Static	Antigen identical
Inventory allocation model 3	Dynamic	Antigen identical

units that are negative for particular minor blood group antigens are only issued to transfusion recipients who are also negative for these antigens. An antigen identical issuing policy results in a realistic average in- and outflow of blood groups, since the distribution of blood groups in the donor and transfusion recipient population was assumed to be the same.

Inventory allocation models 1 and 2 do not consider any variability in the composition of the inventory, as issued RBC units are immediately replaced by RBC units with the same blood groups. To account for randomness in the blood groups supplied, the issued RBC units might also be replaced by RBC units whose blood groups are sampled from the distribution of blood groups in the donor population (dynamic inventory composition). Although an antigen compatible issuing policy would reflect current practice best, it also conflicts with the aim to match all transfusion recipients on every selected antigen. More specifically, an antigen compatible issuing policy could result in shortages of antigen-negative blood groups. Therefore, the dynamic inventory composition was only applied in combination with an antigen identical issuing policy (inventory allocation model 3).

For inventory allocation models 1 and 2, due to their limited complexity, an exact computation of the availability of matched RBC units can be derived by probability calculations. In contrast, for model 3 simulations were required to calculate an average availability of antigen identical RBC units (Section 7.5, Appendix I).

7.2.3　Maximum proportion of alloimmunization prevented

Completely matched RBC units would eliminate all transfusion-induced alloimmunization. However, as inventory sizes are finite, a perfect match on all antigens is often impossible in practice. A match on a smaller number of antigens, in contrast, is feasible, but might lead to alloimmunization against antigens excluded from the matching strategy. To maximize the proportion of alloimmunization prevented, a dynamic programming approach is used to find an optimal sequence for antigen exclusion.

The dynamic programming approach decomposes all possible sequences containing up to twelve minor RBC antigens ($\sim 10^9$) into a small number of levels and connected states ($\sim 10^4$) (Figure 7.1). This reduces computational complexity
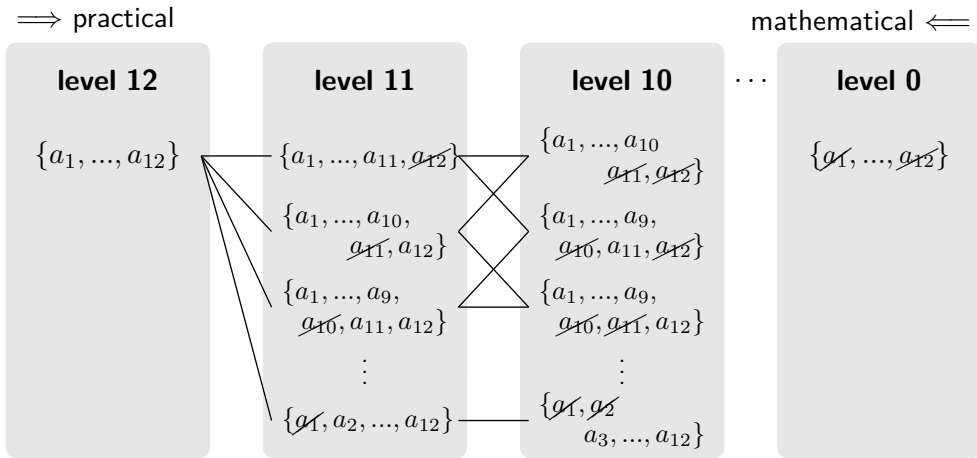
Figure 7.1: State transition diagram, which forms the basis for the dynamic programming approach. It consists of 4,096 different states (i.e., matching strategies) connected by 24,576 arcs. It decomposes all possible sequences ($\sim 10^9$) in which up to twelve minor RBC antigens can be either included in ($\Longleftarrow$ mathematical) or excluded from ($\Longrightarrow$ practical) a matching strategy.

drastically [87]. Each state represents a unique matching strategy $\mathcal{M}$, consisting (next to antigens A, B, D) of ℓ minor RBC antigens:

$$\mathcal{M} = \{A, B, D\} \cup \{\mu_1, ..., \mu_\ell\}. \tag{7.1}$$

As antigens A, B, and D are always included and therefore non-distinctive, we only refer to the set minor RBC antigens, which is denoted by $\mathcal{M}_\ell = \{\mu_1, ..., \mu_\ell\}$.

As illustrated in Figure 7.1, matching strategies can be ordered into a state transition diagram, where each matching strategy $\mathcal{M}_\ell$ is assigned to a level ℓ equal to the number of minor RBC antigens considered. Moreover, matching strategies between two consecutive levels are connected if the matching strategy at level ℓ contains exactly one antigen that is not part of the matching strategy at level $\ell-1$. All sequences in which antigens could be excluded from a matching strategy can be found by traversing the state transition diagram from left to right ($\Longrightarrow$). A realistic approach in practice is to exclude antigens whenever the inventory is too small to allow matching for all antigens considered. However, to find the order of antigens to be excluded maximizing the proportion of alloimmunization prevented, the state transition diagram has to be solved in the opposite direction, i.e., from right to left ($\Longleftarrow$).

The maximum proportion of alloimmunization prevented, given a particular matching strategy $\mathcal{M}_\ell$, was calculated iteratively by combining the maximum proportion of alloimmunization prevented at the previous level plus the effect of adding one additional antigen at level ℓ. More specifically, for each level ℓ and for each
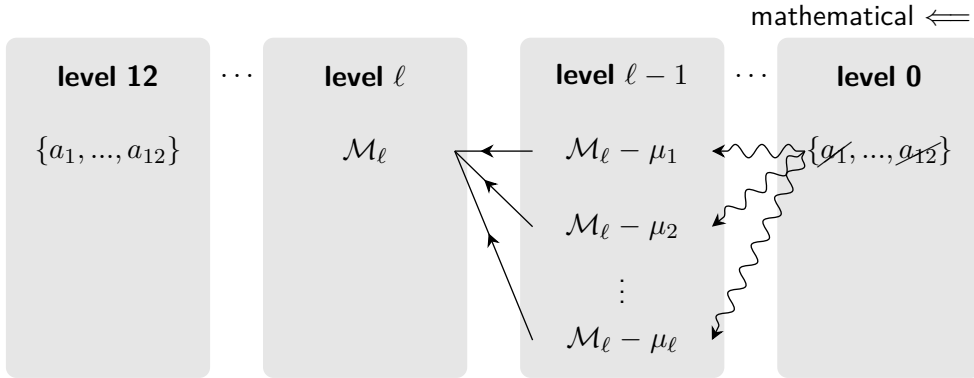
Figure 7.2: State transition diagram. It illustrates how the maximum proportion of alloimmunization prevented for matching strategy $\mathcal{M}_\ell = \{\mu_1, ..., \mu_\ell\}$ is computed.

matching strategy $\mathcal{M}_\ell$ at level ℓ, the maximum proportion of alloimmunization prevented was calculated by considering all paths up to level $\ell - 1$ (which have already been optimized) and selecting the path leading to the highest proportion of alloimmunization prevented (Figure 7.2). Note that this is not the same as optimizing the effect of adding one additional antigen to the optimal sequence obtained so far, as the order in which antigens are added might change as all paths for the new matching strategy are reconsidered.

In essence, given that level $\ell - 1$ has been solved, yet a separate optimization problem exists in each state $\mathcal{M}_\ell = \{\mu_1, ..., \mu_\ell\}$ at level ℓ, which maximizes the proportion of alloimmunization prevented by matching strategy $\mathcal{M}_\ell$, i.e.,

$$V(\mathcal{M}_\ell) = \max_{k=1,...,\ell} \left\{ r(\mathcal{M}_\ell - \mu_k \to \mathcal{M}_\ell) + V(\mathcal{M}_\ell - \mu_k) \right\} \tag{7.2}$$

where

- $V(\mathcal{M}_\ell)$ represents the maximum proportion of alloimmunization prevented by matching strategy $\mathcal{M}_\ell$ (consisting ℓ minor RBC antigens in addition to A, B and D),

- $V(\mathcal{M}_\ell - \mu_k)$ represents the maximum proportion of alloimmunization prevented by matching strategy $\mathcal{M}_\ell - \mu_k$ (consisting $\ell - 1$ minor RBC antigens in addition to A, B and D),

- $r(\mathcal{M}_\ell - \mu_k \to \mathcal{M}_\ell)$ represents the effect of adding antigen μ_k to matching strategy $\mathcal{M}_\ell - \mu_k$.

The value of each matching strategy $\mathcal{M}_\ell$ can be computed by one-and-the-same recursive formula, which is used repeatedly. The value of $r(\mathcal{M}_\ell - \mu_k \to \mathcal{M}_\ell)$

is not simple an incremental effect of one additional antigen, but depends on the antigens that were already included in matching strategy $\mathcal{M}_\ell - \mu_k$ (Appendix 7.5).

In short, the dynamic programming approach determines the optimal sequence in which antigens have to be included, such that the proportion of alloimmunization prevented is maximized. Here, the position of an antigen indicates the importance of matching on this specific antigen. Hence, antigens that are added at a lower level will prevent more alloimmunization than antigens that are added at a higher level. From a practical perspective, in fact, the reverse sequence must be thought of. It can be interpreted as the optimal sequence in which antigens have to be excluded from a matching strategy when completely matched RBC units cannot be found in inventory.

7.2.4 Scenarios analyzed

In this study the frequencies of fifteen antigens (and cognate antibodies) in Caucasians were taken into consideration: A, B, D, C, c, E, e, K, Fy^a, Fy^b, Jk^a, Jk^b, M, S, and s. Computations are performed for:

- matching strategies from three (A, B, D) up to fifteen antigens,

- three inventory allocation models, which differed in inventory composition (static / dynamic) and issuing policy applied (compatible / identical),

- inventory sizes of a small hospital (60 units), a large hospital (120 units), an academic hospital (250 units), and a distribution centre (1000 units) (these numbers are chosen according to inventory sizes in the Netherlands),

- requests for 1, 2, 3, 5, and 10 units, compatible or identical with the blood group of a single transfusion recipient.

7.3 Results

7.3.1 Proportion of alloimmunization prevented

Computations using inventory allocation model 1 show that, when two RBC units with a particular blood group are requested, the proportion of alloimmunization prevented equals 0.90, 0.95, 0.98, and 1.00 for inventory size of 60, 120, 250, and 1000 units, respectively (Figure 7.3a-d). For inventory allocation model 2, these proportions decrease to 0.72, 0.81, 0.88, and 0.95 (Figure 7.3e-h), and for inventory allocation model 3 to 0.46, 0.63, 0.78, and 0.94 (Figure 7.3i-l). Moreover, when the number RBC units requested increases, these proportions further decrease, where magnitude of the decline strongly depends on the inventory size n.

An important similarity between the graphs in Figure 7.3, is that most of the gain already can be obtained by matching on the two to four most important
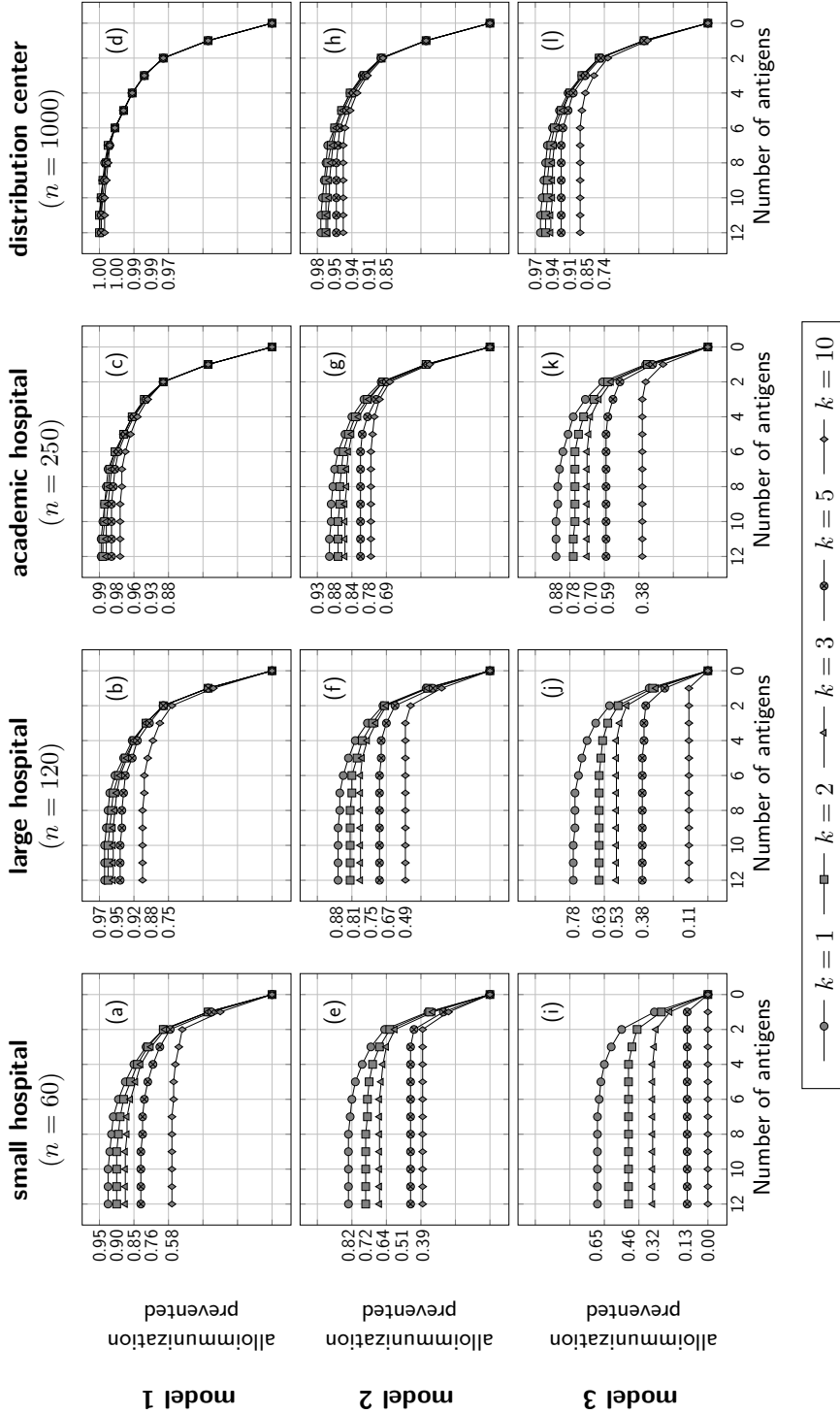
Figure 7.3: Maximum proportion of alloimmunization prevented for $n = 60, 120, 250, 1000$ (inventory size); $k = 1, 2, 3, 5, 10$ (number of units requested for a single transfusion recipient); and inventory allocation models 1, 2, and 3. The increment between two consecutive points on a curve indicates the contribution of an additional antigen on the total proportion of alloimmunization prevented.

minor RBC antigens (i.e., E, K, Jkª, and c), together responsible for 0.63 up to 0.80 of the alloantibodies formed by transfusion recipients (Table 7.1). For example, for inventory allocation model 3 and an inventory size of 60 units, the proportion of alloimmunization prevented, equals 0.60 (for 1 unit) / 0.46 (for 2 units), when matching on these four antigens. These proportions increase to 0.70 / 0.61, 0.78 / 0.72, and 0.80 / 0.80 for inventory sizes of 120, 250, and 1000 units respectively. Moreover, even when 5 up to 10 units are requested from a distribution center, a match on the four most important antigens prevents 0.71 up to 0.78 of alloimmunization events.

7.3.2 Optimal order for antigens exclusion

The increment between two consecutive points on a curve (Figure 7.3) indicates the contribution of the addition of a single antigen on the total proportion of alloimmunization prevented. As these proportions are computed sequentially, the contribution of a particular antigen does not only depend on the proportion of antibodies directed against this antigen (Table 7.1), but also on the antigens that are already matched. This has a substantial impact on the likelihood that a transfusion recipient can be matched for this additional antigen, since it might still be excluded in case insufficient matched RBC units are available from inventory.

The optimal matching strategy at each matching level provides an order in which antigens have to be added. The optimal matching strategy can vary per scenario (inventory size and number of RBC units requested) and was determined for 20 different scenarios (four inventory sizes, five numbers of RBC units requested). Figure 7.4 shows the frequency that antigens occur at a particular position in the
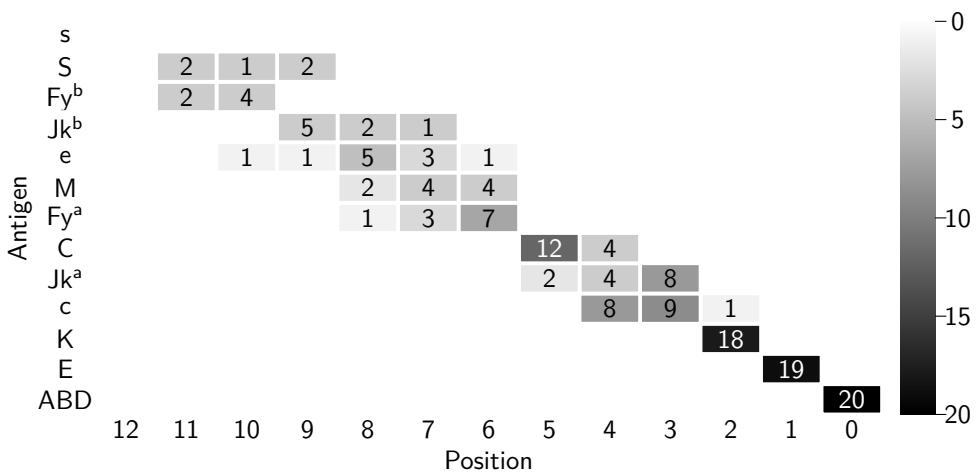
Figure 7.4: Distribution of the optimal positions per antigens within the matching strategy over all 20 scenarios considered. The grey-scale corresponds to the frequency indicated.
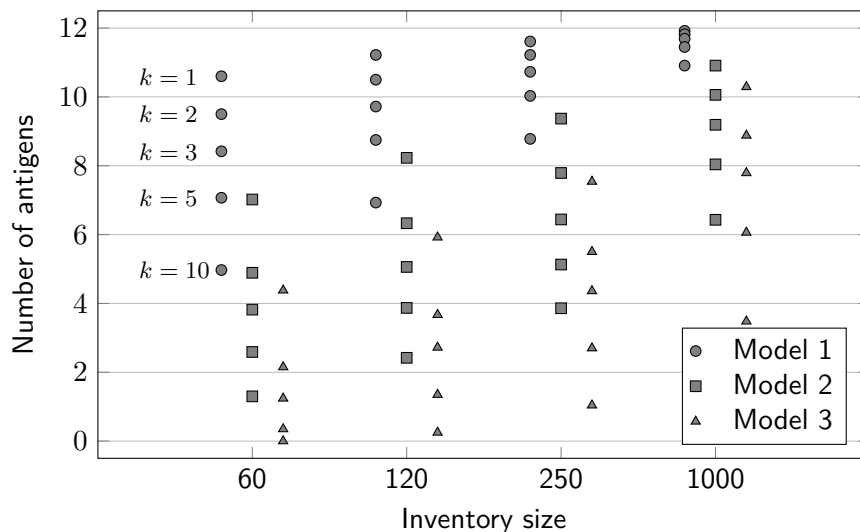
Figure 7.5: Average number of minor RBC antigens for which a transfusion recipient was matched for $n = 60, 120, 250, 1000$ (inventory size); $k = 1, 2, 3, 5, 10$ (number of units requested for a single transfusion recipient); and inventory allocation models 1, 2, and 3.

sequence of matched antigens for inventory allocation model 3. This figure shows that the antigen order is hardly affected by the number of units requested or the size of the inventory. For most scenarios antigen E is added first to the matching strategy, followed by K, c and Jkᵃ as the majority of the antibodies are directed against these antigens.

Note that not all rows and columns add up to 20. This is because in some situations the likelihood of an increase in the prevention of antibody formation becomes near negligible ($< 10^{-3}$), in which case no further antigens were added to the matching strategy.

7.3.3   Number of minor RBC antigens matched

As antigens might be excluded when the inventory contains an insufficient number of matched RBC units, some transfusion recipients are not matched for all twelve minor RBC antigens consider, but only for a limited number of antigens. Figure 7.5 shows the average number of minor RBC antigens for which a transfusion recipient was matched for each of the three inventory allocation models. A complete representation of the underlying distributions can be found in Supplemental Table 7.5 (Section 7.5).

7.4 Discussion

When genotyping of donors and transfusion recipients becomes affordable in the near future, comprehensive preventive matching might become possible. Up until now, the effect of preventive matching on the rate of alloimmunization has never been determined. In this study we applied mathematical techniques to calculate the proportion of transfusion-induced alloimmunization that might be prevented when both donors and transfusion recipients are comprehensively typed. In the situation where all donors and transfusion recipients are typed for the fifteen most clinically relevant antigens (responsible for the far majority of the immunizations encountered in practice) and all transfusion recipients are optimally matched, this proportion was 94% when two RBC units are requested from an inventory of 1000 RBC units. When preventive matching is applied for a limited number of patients and compatible matching may be applied (inventory allocation model 1), even with smaller inventories of 120 RBC units still 95% of alloimmunization events can be prevented.

An optimal sequence for antigen matching was determined for different inventory sizes and number of units requested with a particular blood group. This sequence does not only depend on the immunogenicity of the antigens, but also by the availability of requested blood groups. As a result, for each model and inventory size a slightly different optimal antigen sequence was obtained. In most models E and K were on the first 2 positions, followed by either c or Jk^a, C, or Fy^a, and eventually by M, S, Jk^b, Fy^b, or e. We used the immunogenicity of the blood group antigens as determined in a prospective study in an incident new-user cohort study among patients transfused in the Netherlands [36]. In this study patients who received other than routinely matched (A, B, D) units were excluded. Although this study is the largest study to date (i.e., 472 immunizations in 21,512 recipients), the immunogenicity values are still approximations. The optimal antigen sequence found in our study might therefore change when more accurate or different estimates for the immunogenicity of antigens become available.

In practice, the majority of transfusion recipients still receive two RBC units per transfusion episode [12]. Therefore, we explicitly discuss the results that were obtained for a request of two units ($k = 2$) and its implications for transfusion practice. Three different models were tested for this situation. Inventory allocation model 1 provides estimates in case a limited number of high risk transfusion recipients receive extensively matched RBC units. The proportion of alloimmunization that can be prevented against selected antigens equals 100% for a distribution centre with 1000 units in stock. However, even if compatible units are selected from a large hospital with only 120 RBC units in stock, the proportion of alloimmunization preventable would still be 95%. This latter result implies that at present 95% of immunizations can be prevented in selected high risk patients.

To model the effects of a preventive matching strategy that is applied for

all transfusion recipients, the inventory allocation strategy had to be adapted to prevent inventory depletion. Therefore, with model 2, first the effect of saving homozygous RBC units for patients that are also homozygous for the respective antigens (antigen identical issuing policy) was tested. As expected, this strategy reduces the level of alloimmunization prevention, but still 95% of antibodies against selected antigens can be prevented if two RBC units are issued from a distribution centre and 72% when issuing from a small hospital. Next, the dynamic inventory composition with an identical issuing policy (inventory allocation model 3) was assessed. The proportion of alloimmunization prevented is more than 91% if three or less RBC units are issued from an inventory of 1000 RBC units. But also a considerable proportion of alloimmunization can be prevented from smaller stocks, ranging from 70% to 32% from inventories with 250 to 60 RBC units.

Our results clearly demonstrate that it is feasible to prevent the majority of transfusion-induced antibodies when both donors and transfusion recipients are comprehensively typed and the inventory from which matched RBC units are issued is sufficiently large. In practice, however, the inventory of RBC units in hospitals is limited, indicating that logistics and inventory management will have to be redesigned to achieve the theoretical gain of preventive matching predicted. On the other hand, the majority of all transfusions are elective [37, 112], which implies that matched blood products are not required instantly. Hence, instead of putting unassigned stock in hospital inventories, it might be worthwhile investigating the feasibility of installing a combined stock for several hospitals as this would increase the overall availability of extensively matched RBC units. However, our results demonstrate that even with the current organization of the blood supply a substantial amount of antibody formation may be prevented when extensive preventive matching is implemented.

Limitations of our study concern first of all the fixed inventory sizes. Once RBC units are issued from inventory these are immediately replaced. In practice however, the total number of RBC units in inventory will fluctuate which will affect the likelihood that a particular RBC unit will be available. Second, we assumed that the number of RBC units requested with a particular blood group (k) was fixed. Especially for (relatively) large k and a dynamic inventory composition this can destabilize the inventory as k RBC units with antigen identical blood groups are extracted from inventory per request, while the blood groups of the new RBC units are randomly selected. The likelihood of finding a set of matching RBC units for a single requests for 10 units will therefore be higher than the likelihood calculated, as there will be a mix of the number of RBC units requested. Third, when a match on all fifteen antigen cannot be found, antigens are excluded until an optimal match on a reduced number of antigens is found. The antigens that are ignored are not revisited, so it might be that a match on more antigens was possible.

Besides the availability of matched RBC units, the blood group of the transfu-

sion recipient has to be known to apply preventive matching strategies. Due to the expansion of genetic diagnostics and technical improvements and cost reduction of genotyping platforms, the number of transfusion recipients whose exomes have been fully sequenced will increase and algorithms are available that can extract the blood group of a donor from exome or whole genome sequencing data [69]. Genotyping platforms that can rapidly and accurately determine the most important blood groups are available [122]. However, our study also demonstrates that the majority of alloimmunizations can already be prevented when recipients are only typed for one to four additional antigens, which can also be achieved by serologic typing methods.

The results presented in this study were obtained by using blood group frequencies and estimates of immunogenicity for a Caucasian population. However, our models are generic, and can be applied to any designated donor and patient population with other blood group frequencies and/or immunogenicity data to find a tailored optimal sequence for antigen matching.

In conclusion, we have demonstrated that with present inventory sizes and the proposed sequence for antigen matching, most transfusion-induced alloimmunization events in Caucasian recipients can be prevented when both donors and recipients are comprehensively typed. This means that preventive matching is a strategy that might deserve serious consideration for practical implementation once comprehensive (geno)typing of donors and patients becomes affordable. Reduced alloimmunization events will not only lower the costs associated with antibody identification, but also improve patient health by decreasing the risk of (alloantibody-induced) hemolytic transfusion reactions and HDFN. Our study shows that alloimmunization is not an inevitable risk of blood transfusion, but that it (in analogy to infectious risks) can be prevented. With the availability of blood group genotyping platforms transfusion policy can be changed to the prevention of alloimmunization by matching.

7.5   Appendix I

Before the mathematical description of the three inventory allocation models and the dynamic programming approach is provided, we first present some details about the mathematical notation of haematological concepts (see also Chapter 4). Table 7.3 provides an overview of the different sets, parameters, and variables used in this chapter.

Table 7.3: Overview of the different sets, parameters, and variables used in this section.

Sets	
$\mathcal{P}$	Power set ($\mathcal{P}(\mathcal{S})$: the set of all subsets of $\mathcal{S}$, including the empty set and $\mathcal{S}$ itself)
$\mathcal{A}$	Set of RBC antigens (index a)
$\mathcal{M} \subseteq \mathcal{A}$	Matching strategy (index a)
$\mathcal{M}_\ell \subseteq \mathcal{M}$	Set of minor RBC antigens (index μ)
Parameters	
n	Inventory size
k	Number of RBC units requested for a single transfusion recipient
$f_i(\mathcal{M})$, $f_j(\mathcal{M})$	Proportion of donors and recipients with blood groups $i \in \mathcal{P}(\mathcal{M})$ and $j \in \mathcal{P}(\mathcal{M})$ respectively
$g_j(\mathcal{M})$	Proportion of donors having a blood group compatible with blood group $j \in \mathcal{P}(\mathcal{M})$
p_a	Proportion of alloantibodies directed against antigen $a \in \mathcal{A}$
Variables	
$X_j \sim \mathsf{Bin}(n, f_i(\mathcal{M}))$	Number of RBC units in inventory with a blood group identical with blood group $j \in \mathcal{P}(\mathcal{M})$
$Y_j \sim \mathsf{Bin}(n, g_j(\mathcal{M}))$	Number of RBC units in inventory with a blood group compatible with blood group $j \in \mathcal{P}(\mathcal{M})$
$L_j(\mathcal{M}, n, k)$	Likelihood that k matched RBC units with phenotype $j \in \mathcal{P}(\mathcal{M})$ are present in an inventory of size n, when matching strategy $\mathcal{M}$ is applied
$r(\mathcal{M}_\ell - \mu_k \to \mathcal{M}_\ell)$	Incremental proportion of alloimmunization prevented by adding antigen $\mu_k \in \mathcal{M}_\ell$ to matching strategy $\mathcal{M}_\ell - \mu_k$
$V(\mathcal{M}_\ell)$	Maximal proportion of alloimmunization prevented, when matching strategy $\mathcal{M} = \{\mathsf{A}, \mathsf{B}, \mathsf{D}\} \cup \mathcal{M}_\ell$ is applied

7.5.1   Blood group terminology and notation

The blood group of an individual is determined by antigens $a \in \mathcal{A}$ which are either present or absent on the surface of the individual's RBCs, where $\mathcal{A} = \{a_1, ..., a_n\}$ is a finite set of antigens. An individual is called a-positive if antigen a is present on surface of the RBCs and a-negative if antigen a is absent. Due to the dichotomous structure of antigen (i.e., it is either present or absent) the set of blood groups that can be composed from antigens $a \in \mathcal{A}$, denoted by $\mathcal{B}$, is equal to the power set of $\mathcal{A}$.

A matching strategy $\mathcal{M} \subseteq \mathcal{A}$ represents a specific set of antigens for which the blood groups of donors and transfusion recipients have to be matched. As, with respect to this specific matching strategy, only the antigens $a \in \mathcal{M}$ are of interest, the other antigens that do not belong to this set (i.e., $a \in \mathcal{A} \backslash \mathcal{M}$) can be ignored. Therefore, the blood groups of donors and transfusion recipients can be compressed into blood groups that are contained in the set $\mathcal{P}(\mathcal{M})$.

Let the relative frequency of blood group $i \in \mathcal{P}(\mathcal{M})$ in the donor population and the relative frequency of blood group $j \in \mathcal{P}(\mathcal{M})$ in the transfusion recipient population be denoted by $f_i(\mathcal{M})$ and $f_j(\mathcal{M})$ respectively. For the computation of these relative frequencies we refer the reader to Section 4.4.

7.5.2   Inventory allocation models

Inventory allocation models 1 and 2 have both a static inventory composition (i.e., the inventory is a reflection of the distribution of blood groups in the donor population and remains unchanged after issuing RBC units). Therefore, an exact computation of the probability that matched RBC units are present in inventory can be derived.

Let X_j be a random variable representing the number of RBC units in inventory compatible with blood group $j \in \mathcal{P}(\mathcal{M})$ Then X_j is binomially distributed with parameters n and $g_j(\mathcal{M})$, where n denotes the inventory size and

$$g_j(\mathcal{M}) = \sum_{i:i \leq j} f_i(\mathcal{M})$$

represents the probability that an RBC in inventory is compatible with blood group $j \in \mathcal{P}(\mathcal{M})$. The likelihood that k compatible RBC units are present in inventory, when matching strategy $\mathcal{M}$ is applied, equals

$$L_j^1(\mathcal{M}, n, k) = \mathbb{P}\left[X_j \geq k\right].$$

Similarly, let Y_j be a random variable representing the number of RBC units in inventory identical to a requested blood group $j \in \mathcal{P}(\mathcal{M})$. Then Y_j is binomially distributed with parameters n and $f_j(\mathcal{M})$ and the likelihood that k identical RBC

Algorithm 1: Inventory allocation model 3

    Input   : inventory size (n), number of units requested (k), distribution of
                 blood groups in the donor and transfusion recipient population
                 $(f_i(\mathcal{M}),\ f_j(\mathcal{M}))$

    Output: likelihood that matched RBC units are present in inventory
                 $(L_j^3(\mathcal{M}, n, k))$

1 Construct an initial inventory by drawing n blood groups from the donor
   population;
2 Set $t_j = 0$ and $s_j = 0$ for all $j \in \mathcal{P}(\mathcal{M})$;
3 **for** $k = 1$ *to* $10,000$ **do**
4   Draw a blood group $j \in \mathcal{P}(\mathcal{M})$ from the transfusion recipient
       population;
5      $t_j = t_j + 1$;
6   **if** $Y_j \geq k$ **then**
7          $s_j = s_j + 1$;
8          $Y_j = Y_j - k$;
9   Refill the inventory by drawing k blood groups from the donor
           population;
10   **end**
11 **end**
12 return $L_j^3(\mathcal{M}, n, k) = s_j/t_j$;

---

units are present in inventory, when matching strategy $\mathcal{M}$ is applied, equals

$$L_j^2(\mathcal{M}, n, k) = \mathbb{P}\left[Y_j \geq k\right].$$

Inventory allocation model 3 has a dynamic inventory composition. The availability of identical RBC units is calculated by simulating the supply and demand of RBC units, where in each iteration it is checked whether the blood group requested appears at least k times in inventory. For a detailed description, see Algorithm 1.

7.5.3   Dynamic programming approach

To find an optimal order in which antigens $a \in \mathcal{A}\backslash \{A, B, D\}$ have to be added to a matching strategy $\mathcal{M}$ that maximizes the total proportion of alloimmunization prevented a dynamic programming approach is applied. This mathematical approach computes the maximum proportion of alloimmunization prevented for each matching strategy

$$\mathcal{M} = \{A, B, D\} \cup \mathcal{M}_\ell, \qquad \mathcal{M}_\ell \in \mathcal{P}\left(\mathcal{A}\backslash \{A, B, D\}\right),$$

in a recursive manner.

First, a network of matching strategies is constructed, in which each matching strategy $\mathcal{M} = \{A, B, D\} \cup \mathcal{M}_\ell$ is assigned to a level $\ell \in \{0, ..., 12\}$ based upon the number of minor RBC antigens in the matching strategy (i.e., $\ell = |\mathcal{M}_\ell|$). Moreover, matching strategies between two consecutive levels are connected if the matching strategy at the lower level ($\mathcal{M}_\ell - \mu_k$) contains exactly one antigen less than the matching strategy at the upper level ($\mathcal{M}_\ell$), as illustrated in Figure 7.1.

The proportion of alloimmunization prevented when transfusion recipients are only matched for antigens A, B and D (i.e., $V(\mathcal{M}_0)$), is set equal to zero. Next, for any $\ell = 1, 2, ..., 12$ and any matching strategy $\mathcal{M}_\ell$ at level ℓ, the maximum proportion of alloimmunization prevented can be calculated by the following recursive formula:

$$V(\mathcal{M}_\ell) = \max_{k=1,...,\ell} \left\{ r(\mathcal{M}_\ell - \mu_k \to \mathcal{M}_\ell) + V(\mathcal{M}_\ell - \mu_k) \right\},$$

where $r(\mathcal{M}_\ell - \mu_k \to \mathcal{M}_\ell)$ evaluates the effect of adding antigen μ_k to a matching strategy $\mathcal{M}_\ell - \mu_k$ on the total proportion of alloimmunization prevented. More specifically,

$$r(\mathcal{M}_\ell - \mu_k \to \mathcal{M}_\ell) = p_{\mu_k} \sum_{j \in \mathcal{P}(\mathcal{M}_\ell)} w(j, \mu_k) L_j(\mathcal{M}_\ell, n, k), \tag{7.3}$$

with

- p_{μ_k} – proportion of antibodies formed against antigen μ_k,

- $w(j, \mu_k)$ – weight function representing the probability that a blood group $j \in \mathcal{P}(\mathcal{M}_\ell)$ is requested given that it contains antigen $\mu_k \in \mathcal{M}_\ell$, which is defined as

$$w(j, \mu_k) = \begin{cases} \dfrac{f_j}{\displaystyle\sum_{i \in \mathcal{P}(\mathcal{M}_\ell) : \mu_k \in \mathcal{I}} f_i}, & \text{if } \mu_k \in \mathcal{J}, \\ 0, & \text{otherwise,} \end{cases}$$

- $L_j(\mathcal{M}_\ell, n, k)$ – likelihood that k matched RBC units with blood group $j \in \mathcal{P}(\mathcal{M})$ are present in an inventory of size n, when matching strategy $\mathcal{M}_\ell$ is applied.

For a detailed description of the dynamic programming approach, see Algorithm 2.

Algorithm 2: Dynamic programming approach

Input : proportion of alloantibodies directed against antigen $a \in \mathcal{A}$ (p_a), distribution of phenotypes in the transfusion recipient population ($f_j(\mathcal{M})$), likelihood that matched RBC units are present in inventory ($L_j(\mathcal{M}, n, k)$)

Output: Maximal proportion of alloimmunization prevented ($V(\mathcal{M}_\ell)$)

---

1 Set $V(\mathcal{M}_0) = 0$;
2 **for** $\ell = 1, ..., 12$ **do**
3   **for** $\mathcal{M}_\ell \in \mathcal{P}(\mathcal{A} \setminus \{A, B, D\}$ *with* $|\mathcal{M}_\ell| = \ell$ **do**
4   **for** $k = 1, ..., \ell$ **do**
5   Compute
         $r(\mathcal{M}_\ell - \mu_k \rightarrow \mathcal{M}_\ell) = p_{\mu_k} \sum_{j \in \mathcal{P}(\mathcal{M})} w(j, a) L_j(\mathcal{M}, n, k)$;
6   **end**
7   Compute $V(\mathcal{M}_\ell) = \max\limits_{k=1,...,\ell} \{r(\mathcal{M}_\ell - \mu_k \rightarrow \mathcal{M}_\ell) + V(\mathcal{M}_\ell - \mu_k)\}$;
8   **end**
9 **end**
10 return $V(\mathcal{M}_\ell)$ for all $\mathcal{M}_\ell \in \mathcal{P}(\mathcal{A} \setminus \{A, B, D\})$;

7.6 Appendix II

n	k	*	0	1	2	3	4	5	6	7	8	9	10	11	12	average
							Number of minor RBC antigens matched									
	1	0.00	0.00	0.00	0.01	0.01	0.01	0.03	0.05	0.01	0.04	0.06	0.06	0.05	0.66	10.60
	2	0.00	0.00	0.00	0.03	0.03	0.01	0.07	0.08	0.02	0.07	0.09	0.07	0.05	0.48	9.50
60	3	0.01	0.00	0.00	0.04	0.04	0.02	0.10	0.09	0.08	0.10	0.07	0.01	0.05	0.37	8.42
	5	0.05	0.01	0.01	0.05	0.09	0.04	0.12	0.10	0.09	0.10	0.06	0.01	0.04	0.24	7.07
	10	0.14	0.01	0.01	0.15	0.13	0.07	0.11	0.10	0.05	0.07	0.04	0.00	0.02	0.11	4.97
	1	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.03	0.03	0.04	0.05	0.04	0.78	11.24
	2	0.00	0.00	0.00	0.00	0.02	0.01	0.04	0.06	0.02	0.05	0.07	0.07	0.06	0.63	10.52
120	3	0.00	0.00	0.00	0.00	0.03	0.02	0.06	0.08	0.02	0.06	0.08	0.08	0.06	0.52	9.93
	5	0.00	0.00	0.00	0.04	0.04	0.02	0.09	0.10	0.08	0.10	0.08	0.01	0.05	0.39	8.72
	10	0.05	0.01	0.01	0.06	0.09	0.04	0.12	0.10	0.09	0.10	0.06	0.01	0.04	0.23	6.95
	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.01	0.02	0.03	0.03	0.87	11.62
	2	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.03	0.03	0.04	0.05	0.05	0.77	11.24
250	3	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.03	0.05	0.04	0.06	0.06	0.06	0.68	10.89
	5	0.00	0.00	0.00	0.00	0.02	0.01	0.04	0.07	0.02	0.06	0.08	0.08	0.06	0.56	10.20
	10	0.00	0.00	0.00	0.04	0.04	0.02	0.10	0.10	0.08	0.10	0.08	0.01	0.05	0.38	8.74
	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.97	11.92
	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.02	0.02	0.93	11.82
1000	3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.03	0.03	0.89	11.72
	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.02	0.03	0.04	0.04	0.83	11.51
	10	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.03	0.04	0.04	0.06	0.06	0.06	0.71	11.03

(a) Inventory allocation model 1

n	k	*	0	1	2	3	4	5	6	7	8	9	10	11	12	average
							Number of minor RBC antigens matched									
	1	0.02	0.02	0.04	0.06	0.06	0.02	0.04	0.14	0.12	0.12	0.13	0.06	0.03	0.14	7.02
	2	0.04	0.04	0.07	0.11	0.07	0.09	0.20	0.01	0.18	0.07	0.06	0.02	0.01	0.02	4.89
60	3	0.08	0.05	0.09	0.05	0.10	0.20	0.21	0.15	0.04	0.03	0.00	0.01	0.00	0.00	3.82
	5	0.17	0.07	0.09	0.10	0.15	0.24	0.14	0.02	0.01	0.00	0.00	0.00	0.00	0.00	2.59
	10	0.27	0.07	0.25	0.27	0.05	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.30
	1	0.01	0.01	0.02	0.04	0.04	0.01	0.03	0.10	0.11	0.11	0.15	0.08	0.05	0.24	8.23
	2	0.02	0.02	0.05	0.07	0.05	0.06	0.15	0.02	0.20	0.11	0.11	0.05	0.03	0.06	6.33
120	3	0.03	0.03	0.07	0.04	0.05	0.14	0.20	0.22	0.09	0.08	0.00	0.03	0.01	0.02	5.06
	5	0.05	0.05	0.09	0.04	0.08	0.26	0.26	0.10	0.05	0.01	0.01	0.00	0.00	0.00	3.87
	10	0.19	0.09	0.06	0.10	0.19	0.28	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.00	2.42
	1	0.00	0.00	0.01	0.02	0.03	0.01	0.01	0.06	0.08	0.08	0.14	0.09	0.06	0.39	9.37
	2	0.01	0.01	0.02	0.04	0.03	0.05	0.10	0.02	0.15	0.13	0.14	0.09	0.05	0.16	7.79
250	3	0.01	0.02	0.04	0.03	0.04	0.08	0.13	0.20	0.14	0.14	0.01	0.07	0.03	0.07	6.44
	5	0.02	0.03	0.06	0.04	0.04	0.11	0.24	0.20	0.14	0.05	0.03	0.00	0.01	0.01	5.13
	10	0.05	0.07	0.07	0.03	0.06	0.33	0.25	0.08	0.05	0.02	0.00	0.00	0.00	0.00	3.86
	1	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.04	0.04	0.08	0.06	0.05	0.69	10.91
	2	0.00	0.00	0.00	0.01	0.01	0.02	0.04	0.01	0.07	0.09	0.09	0.09	0.07	0.51	10.06
1000	3	0.00	0.00	0.01	0.00	0.01	0.04	0.06	0.09	0.11	0.11	0.02	0.10	0.07	0.38	9.19
	5	0.00	0.01	0.01	0.01	0.02	0.06	0.07	0.12	0.12	0.17	0.12	0.01	0.07	0.21	8.04
	10	0.01	0.02	0.02	0.02	0.05	0.08	0.10	0.18	0.19	0.16	0.08	0.04	0.00	0.05	6.43

(b) Inventory allocation model 2

| n | k | * | \multicolumn{13}{c}{Number of minor RBC antigens matched} | average |
			0	1	2	3	4	5	6	7	8	9	10	11	12	
	1	0.10	0.06	0.12	0.05	0.11	0.07	0.10	0.12	0.12	0.04	0.02	0.03	0.01	0.05	4.38
	2	0.19	0.10	0.17	0.16	0.14	0.07	0.12	0.04	0.01	0.00	0.00	0.00	0.00	0.00	2.15
60	3	0.25	0.16	0.25	0.17	0.06	0.09	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.24
	5	0.39	0.29	0.29	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35
	10	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	1	0.04	0.04	0.04	0.06	0.12	0.03	0.11	0.14	0.14	0.06	0.06	0.06	0.02	0.09	5.83
	2	0.07	0.07	0.10	0.11	0.23	0.06	0.10	0.18	0.04	0.02	0.01	0.00	0.00	0.00	3.46
120	3	0.12	0.09	0.14	0.14	0.31	0.13	0.01	0.05	0.01	0.00	0.00	0.00	0.00	0.00	2.32
	5	0.20	0.13	0.25	0.19	0.19	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.37
	10	0.38	0.33	0.27	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33
	1	0.02	0.00	0.04	0.07	0.01	0.09	0.01	0.12	0.16	0.10	0.12	0.05	0.02	0.19	7.40
	2	0.03	0.04	0.03	0.13	0.04	0.15	0.20	0.06	0.16	0.08	0.04	0.02	0.01	0.01	4.92
250	3	0.04	0.06	0.05	0.10	0.18	0.18	0.04	0.23	0.08	0.02	0.01	0.00	0.00	0.00	4.00
	5	0.08	0.07	0.12	0.14	0.31	0.17	0.06	0.03	0.01	0.00	0.00	0.00	0.00	0.00	2.63
	10	0.20	0.11	0.28	0.25	0.12	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.34
	1	0.00	0.00	0.00	0.01	0.01	0.02	0.00	0.04	0.06	0.07	0.07	0.15	0.05	0.51	10.25
	2	0.01	0.00	0.00	0.01	0.02	0.02	0.02	0.07	0.16	0.13	0.21	0.10	0.05	0.20	8.70
1000	3	0.01	0.00	0.00	0.02	0.02	0.06	0.03	0.11	0.24	0.16	0.12	0.13	0.03	0.06	7.58
	5	0.01	0.00	0.02	0.08	0.01	0.12	0.22	0.06	0.25	0.12	0.05	0.04	0.01	0.01	5.88
	10	0.03	0.03	0.06	0.07	0.16	0.23	0.05	0.28	0.07	0.01	0.01	0.00	0.00	0.00	4.22

(c) Inventory allocation model 3

Table 7.4: Number of antigens matched for $n = 60, 120, 250, 1000$ (inventory size), $k = 1, 2, 3, 5, 10$ (number of units requested for a single transfusion recipients). *For some combinations of n and k a match on only antigen A, B, and D was not possible.

| n | k | Position |
		1	2	3	4	5	6	7	8	9	10	11	12
	1	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
	2	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b		
60	3	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b		
	5	E	K	c	Jk^a	C	Fy^a	M	S	Jk^b			
	10	E	K	Jk^a	c	C	Fy^a	S					
	1	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
	2	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b	e	
120	3	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b		
	5	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b		
	10	E	K	c	Jk^a	C	Fy^a	M	S	Jk^b			
	1	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
	2	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
250	3	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b	e	
	5	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b		
	10	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b		
	1	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
	2	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
1000	3	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
	5	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
	10	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b	Fy^b	e	

(a) Inventory allocation model 1

n	k	1	2	3	4	5	6	7	8	9	10	11	12
							Position						
60	1	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b			
	2	E	K	c	Jk^a	C	Fy^a	S					
	3	E	K	c	C	Jk^a	Fy^a	S					
	5	E	K	c	C								
	10	E	K										
120	1	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b Fy^b	e		
	2	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b			
	3	E	K	c	Jk^a	C	Fy^a	S					
	5	E	K	c	C	Fy^a	S						
	10	E	K	c	C	S							
250	1	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	e	Fy^b	
	2	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b			
	3	E	K	Jk^a	c	C	Fy^a	M	S	Jk^b			
	5	E	K	Jk^a	c	C	Fy^a	S					
	10	E	K	c	C	Jk^a	Fy^a						
1000	1	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
	2	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	Fy^b	e	
	3	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	e	Fy^b	
	5	E	K	Jk^a	c	Fy^a	C	M	S	Jk^b	e		
	10	E	K	Jk^a	c	Fy^a	C	S	M	Jk^b			

(b) Inventory allocation model 2

n	k	1	2	3	4	5	6	7	8	9	10	11	12
							Position						
60	1	E	K	c	C	Jk^a	M	Fy^a	e	Jk^b	Fy^b		
	2	E	K	c	C	Jk^a	e						
	3	E	K	c	C								
	5	E	K										
	10												
120	1	E	K	Jk^a	c	C	M	Fy^a	e	Jk^b	Fy^b	S	
	2	E	K	c	Jk^a	C	M	e	Fy^a				
	3	E	K	c	Jk^a	C							
	5	E	K	c	C								
	10	E	c										
250	1	E	K	Jk^a	c	C	Fy^a	Jk^b	M	S	Fy^b		
	2	E	K	Jk^a	c	C	Fy^a	M	e	Jk^b			
	3	E	K	c	Jk^a	C	Fy^a	e	M				
	5	E	K	c	Jk^a	C							
	10	E	K	c									
1000	1	E	K	Jk^a	c	C	Fy^a	M	Jk^b	S	e	Fy^b	
	2	E	K	Jk^a	c	C	Fy^a	M	Jk^b	e	S	Fy^b	
	3	E	K	Jk^a	c	C	M	Fy^a	e	Jk^b	Fy^b	S	
	5	E	K	Jk^a	c	C	Fy^a	M	e	Jk^b			
	10	E	K	Jk^a	c	C	Fy^a	e					

(c) Inventory allocation model 3

Table 7.5: Optimal order in which antigens have to be added to a matching strategy for $n = 60, 120, 250, 1000$ (inventory size), $k = 1, 2, 3, 5, 10$ (number of units requested for a single transfusion recipients).

Conclusion and outlook

Healthcare is moving towards precision medicine, which offers a more in-depth understanding of human physiology using genetic insights and technological advances. This is pivotal in alleviating unnecessary suffering due to unintended side effects, which result from the currently applied one-size-fits-all approach [61]. New genetic insights and technological advances also shape the field of transfusion medicine. Inexpensive molecular typing paired with powerful bioinformatics has enabled mass-scale red blood cell genotyping of donors and transfusion recipients [65]. These will likely replace serology-based methods and increase the availability of extended antigen matched red blood cell units [15, 65]. Over the past years, extended antigen matching has become standard practice for some transfusion recipients and has proven to reduce alloimmunization risks. As stated in [15]: "Genotyping has the potential to facilitate the optimization of red cell component provisioning on a mass-scale by improving the efficiency by which the supply of antigen-negative units meets demand". This thesis will evaluate this potential by using methods from the mathematical discipline of Operations Research (OR). To do so, a comprehensively antigen-typed donor and transfusion recipient population is presumed. Accordingly, throughout this thesis, arbitrary blood groups have been considered. For illustrative purposes, ABO, D examples have been included.

A major challenge faced by blood transfusion services is to maintain an adequate donor base, with a sufficiently diverse age distribution and blood group composition [16]. This composition may restrict the ability to select extended matched red blood cell units for transfusion. In fact, an overrepresentation of donors with antigen-negative blood groups is strived for, so that their red blood cells can be matched with a larger proportion of the recipient population. Although the blood group of an individual is normally unknown, relatives are more likely to have blood groups that are similar. Therefore knowledge of the blood groups of active donors can be used to target relatives of donors with desirable (rare) blood groups for recruitment. This has been made practical in Chapter 3 for arbitrary blood groups beyond ABO, D. Several next of kin relations were computed; this is

a novel contribution to the literature.

More generally, in this thesis, several mathematical models are developed for the optimization of blood inventory management and particularly the issuing of available RBC units. As extended blood groups are to be described, a first major challenge was to establish an unambiguous mathematical description to represent arbitrary blood groups, irrespective of the antigens considered. Instead of describing a blood group by indicating the names of the antigens that are present or absent, we represented blood groups by binary vectors. By an ordering principle, this binary description incorporates compatibility between blood groups, implying that a compatibility matrix is no longer required. By representing blood groups by binary vectors and disposing of the compatibility matrix, we have provided a generic representation that can be used to model blood group allocation or matching. Furthermore, this binary representation simplifies the application of advanced mathematical OR techniques for optimization. This has been made apparent in Chapters 5, 6, and 7 and has been demonstrated in Chapters 6 and 7 by practical examples.

An important question addressed in this thesis is how the inventory on hand should be issued, such that all requests can be satisfied with a compatible RBC unit, both directly upon requests and for future demands. Outdating levels should at the same time be minimized as RBC units form a precious good that is voluntarily donated by non-remunerated donors. As a result, the ability to issue highly compatible RBC units eliminates the risk of adverse events as a result of incompatible transfusions. The starting point of issuing RBC units is a matching strategy that determines which antigens should be matched, and hence, which RBC units are expected to be compatible with those of a requested blood product. With an increase in the number of antigens considered in a matching strategy, the level of alloimmunization prevented will increase as well, but it becomes more challenging to find compatible units. The matching strategy chosen, therefore, affects the probability of being able to issue compatible RBC units and the proportion of alloimmunization prevented by matching.

In Chapter 5, the problem of optimization RBC issuing is transformed into a discrete and computational feasibile structure by an MDP description. It incorporates all relevant aspects of donations and requests occurring at arbitrary times, as well as the ageing of RBC units and blood group compatibility. The mathematical model developed allows derivation of an optimal issuing policy for any possible blood group requested and for any possible inventory composition. However, this model is still computational prohibitive, as the number of possible inventory compositions grows rapidly with the number of antigens considered. Nevertheless, by its discrete and MDP structure, there are several possible directions for further research to find an approximate optimal allocation strategy. One of them is to predetermine an issuing policy in some sub-optimal manner and evaluate its quality. This was done in Chapter 6. Other approaches could be to consider a limited num-

ber of states or cluster states with similar structures (e.g., by using approximate dynamic programming and reinforcement learning).

In Chapter 6, a predetermined issuing policy is search for, which aims to balance the age and rareness of the RBC units in inventory. This issuing policy is incorporated into a deterministic optimization model (a minimum cost flow problem). For a given inventory composition, the optimization model determines which RBC units are selected to satisfy requests. Its performance is evaluated by simulation.

When transfusion recipients are matched for just the ABO, D blood groups, no shortage or outdating occurs. When this predetermined strategy is applied for extended blood groups (i.e., up to fourteen different antigens), both shortage or outdating slightly increase. Hence, the majority of requests can be satisfied with an antigen compatible red blood cell unit. The reasons this works well, is that we assumed that 1) the distribution of blood groups in the donor and transfusion recipient population to be identical and 2) the inventories to be sufficiently large, such that in most cases an antigen identical match on all fourteen antigens considered is possible.

In practice, a shortage can easily be resolved. It merely requires that a subgroup of antigens is matched. The antigens that are not matched might elicit an immune response in the transfusion recipient. The probability of such an immune response depends on the immunogenicity of the antigens. Chapter 7 investigates which of the antigens are best not to be matched for in these cases. An optimal matching order could be determined, which maximizes the proportion of alloimmunization prevented. A vast majority of the alloimmunization events could so be prevented by extended matching.

Currently, the selection of RBC units to be issued is a manual process. This will no longer be possible when the number of antigens that have to be matched is extended, as the inventory will rapidly become too complicated for a human to manage. Hence, a system that is able to provide a fast and smart selection of RBC units to be issued is recommended. The models presented in this thesis will stimulate the development of such systems. The work presented in this thesis shows that

- the majority of alloimmunization can be prevented by a matching strategy that is extended with a relatively limited number antigens,

- a practical solution for selecting the most suitable RBC units can be provided based on mathematical optimization.

These results clearly show that blood safety can be improved by merely enhancing the distribution and allocation of the blood products currently available. The cost-effectiveness of the implementation of such a strategy is a topic for further research.

Bibliography

[1] U. Abdulwahab and M.I.M. Wahab. Approximate dynamic programming modeling for a typical blood platelet bank. *Computers & Industrial Engineering*, 78:259–270, 2014.

[2] I.J.B.F. Adan, M.J.A. van Eenige, and J.A.C. Resing. Fitting discrete distributions on the first two moments. *Probability in the Engineering and Informational Sciences*, 9(4):623–632, 1995.

[3] M.P. Atkinson, M.J. Fontaine, L.T. Goodnough, and L.M. Wein. A novel allocation strategy for blood transfusions: investigating the tradeoff between the age and availability of transfused blood. *Transfusion*, 52(1):108–117, 2012.

[4] M. Bakker, J. Riezebos, and R.H. Teunter. Review of inventory systems with deterioration since 2001. *European Journal of Operational Research*, 221(2):275–284, 2012.

[5] T.F. Baskett. James Blundell: the first transfusion of human blood. *Resuscitation*, 52(3):229–233, 2002.

[6] M.P. Bauer, J.C. Wiersum-Osselton, M. Schipperus, J.P. Vandenbroucke, and E. Briët. Clinical predictors of alloimmunization after red blood cell transfusion. *Transfusion*, 47(11):2066–2071, 2007.

[7] J. Beliën and H. Forcé. Supply chain management of blood products: A literature review. *European Journal of Operational Research*, 217(1):1–16, 2012.

[8] S.N. Bernstein. Solution of a mathematical problem connected with the theory of heredity. *The Annals of Mathematical Statistics*, 13(1):53–61, 1942.

[9] D. Bertsimas and J.N. Tsitsiklis. *Introduction to linear optimization*. Athena Scientific, Belmont, CA, 1997.

[10] J.T. Blake, M. Hardy, G. Delage, and G. Myhal. Déjà-vu all over again: using simulation to evaluate the impact of shorter shelf life for red blood cells at Héma-Québec. *Transfusion*, 53(7):1544–1558, 2013.

[11] J. Blundell. Observations on transfusion of blood. with a description of his gravitator. *Lancet*, 2:321–324, 1828.

[12] B.A. Borkent-Raven, M.P. Janssen, C.L. van der Poel, W.P. Schaasberg, G.J. Bonsel, and B.A. van Hout. The proton study: profiles of blood product transfusion recipients in the netherlands. *Vox sanguinis*, 99(1):54–64, 2010.

[13] R.J. Boucherie and N.M. Van Dijk. *Markov decision processes in practice*. Springer, Cham, Switzerland, 2017.

[14] D.C. Brousseau, J.A. Panepinto, M. Nimmer, and R.G. Hoffmann. The number of people with sickle-cell disease in the United States: national and state estimates. *American Journal of Hematology*, 85(1):77–78, 2010.

[15] J.H. Carter and W.A. Flegel. Red cell transfusions in the genomics era. *Seminars in Hematology*, 56(4):236–240, 2019.

[16] M.C. Carter, J. Wilson, G.S. Redpath, P. Hayes, and C. Mitchell. Donor recruitment in the 21st century: challenges and lessons learned in the first decade. *Transfusion and Apheresis Science*, 45(1):31–43, 2011.

[17] J.M. Casas, M. Ladra, B.A. Omirov, and R. Turdibaev. On the algebraic properties of the human ABO-blood group inheritance pattern. *The ANZIAM Journal*, 58(1):78–95, 2016.

[18] CBO. CBO blood transfusion guideline, 2011. `http://www.isbtweb.org/file admin/user_upload/blood-transfusion-guideline.pdf` [Accessed: 17 January 2020].

[19] S.T. Chou, R.I. Liem, and A.A. Thompson. Challenges of alloimmunization in patients with haemoglobinopathies. *British Journal of Haematology*, 159 (4):394–404, 2012.

[20] S.T. Chou, J.M. Flanagan, S. Vege, N.L.C. Luban, R.C. Brown, R.E. Ware, and C.M. Westhoff. Whole-exome sequencing for RH genotyping and alloimmunization risk in children with sickle cell anemia. *Blood Advances*, 1 (18):1414–1422, 2017.

[21] S.T. Chou, P. Evans, S. Vege, S.L. Coleman, D.F. Friedman, M. Keller, and C.M. Westhoff. RH genotype matching for transfusion support in sickle cell disease. *Blood*, 132(11):1198–1207, 2018.

[22] Y.T. Chung and F. Erhun. Designing supply contracts for perishable goods with two periods of shelf life. *IIE Transactions*, 45(1):53–67, 2013.

[23] I. Civelek, I. Karaesmen, and A. Scheller-Wolf. Blood platelet inventory management with protection levels. *European Journal of Operational Research*, 243(3):826–838, 2015.

[24] R.R.A. Coombs, A.E. Mourant, and R.R. Race. A new test for the detection of weak and "incomplete" Rh agglutinins. *British Journal of Experimental Pathology*, 26(4):255–266, 1945.

[25] R.R.A. Coombs, A.E. Mourant, and R.R. Race. In-vivo isosensitisation of red cells in babies with haemolytic disease. *The Lancet*, 247(6391):264–266, 1946.

[26] R.J. Davey. Recruiting blood donors: challenges and opportunities. *Transfusion*, 44(4):597–600, 2004.

[27] M. de Haas, F.F. Thurik, J.M. Koelewijn, and C.E. van der Schoot. Haemolytic disease of the fetus and newborn. *Vox Sanguinis*, 109(2):99–113, 2015.

[28] W.L.A.M. de Kort, M. Janssen, N. Kortbeek, N. Jansen, J. van der Wal, and N.M. van Dijk. Platelet pool inventory management: theory meets practice. *Transfusion*, 51(11):2295–2303, 2011.

[29] M. Dillon, F. Oliveira, and B. Abbasi. A two-stage stochastic programming model for inventory management in the blood supply chain. *International Journal of Production Economics*, 187:27–41, 2017.

[30] Q. Duan and T.W. Liao. A new age-based replenishment policy for supply chain inventory optimization of highly perishable products. *International Journal of Production Economics*, 145(2):658–671, 2013.

[31] Q. Duan and T.W. Liao. Optimization of blood supply chain with shortened shelf lives and ABO compatibility. *International Journal of Production Economics*, 153:113–129, 2014.

[32] G. Edgren, H. Hjalgrim, T.N. Tran, K. Rostgaard, A. Shanwell, K. Titlestad, L. Jakobsson, G. Gridley, L. Wideroff, C. Jersild, J. Adami, M. Melbye, M. Reilly, and O. Nyrén. A population-based binational register for monitoring long-term outcome and possible disease concordance among blood donors and recipients. *Vox Sanguinis*, 91(4):316–323, 2006.

[33] G. Edgren, K. Rostgaard, S.K. Vasan, A. Wikman, R. Norda, O.B. Pedersen, C. Erikstrup, K.R. Nielsen, K. Titlestad, H. Ullum, M. Melbye, O. Nyrén, and

H. Hjalgrim. The new Scandinavian Donations and Transfusions database (SCANDAT2): a blood safety resource with added versatility. *Transfusion*, 55(7):1600–1606, 2015.

[34] M.O. Elebute, L. Choo, A. Mora, C. MacRury, C. Llewelyn, S. Purohit, V. Hicks, C. Casey, M. Malfroy, A. Deary, T. Reed, S. Meredith, L. Manson, and L.M. Williamson. Transfusion of prion-filtered red cells does not increase the rate of alloimmunization or transfusion reactions in patients: results of the UK trial of prion-filtered versus standard red cells in surgical patients (PRISM A). *British Journal of Haematology*, 160(5):701–708, 2013.

[35] H. Ensafian, S. Yaghoubi, and M.M. Yazdi. Raising quality and safety of platelet transfusion services in a patient-based integrated supply chain under uncertainty. *Computers & Chemical Engineering*, 106:355–372, 2017.

[36] D. Evers, R.A. Middelburg, M. de Haas, S. Zalpuri, K.M.K. de Vooght, D. van de Kerkhof, O. Visser, N.C. Péquériaux, F. Hudig, H. Schonewille, J.J. Zwaginga, and van der Bom J.G. Red-blood-cell alloimmunisation in relation to antigens' exposure and their immunogenicity: a cohort study. *The Lancet Haematology*, 3(6):e284–e292, 2016.

[37] A.M. Fillet, M. Desmarets, S. Assari, J.F. Quaranta, A. François, A. Pugin, A. Schuhmacher, B. Lassale, E. Monnet, P. Cabre, D. Legrand, Binda D., and Djoudi R. Blood products use in france: a nationwide cross-sectional survey. *Transfusion*, 56(12):3033–3041, 2016.

[38] W.A. Flegel. Red cell alloimmunisation: incidence and prevention. *The Lancet Haematology*, 3(6):e260–e261, 2016.

[39] W.A. Flegel, C. Natanson, and H.G. Klein. Does prolonged storage of red blood cells cause harm? *British Journal of Haematology*, 165(1):3–16, 2014.

[40] W.A. Flegel, J.L. Gottschall, and G.A. Denomme. Integration of red cell genotyping into the blood supply chain: a population-based study. *The Lancet Haematology*, 2(7):e282–e289, 2015.

[41] M.J. Fontaine, Y.T. Chung, W.M. Rogers, H.D. Sussmann, P. Quach, S.A. Galel, L.T. Goodnough, and F. Erhun. Improving platelet supply chains through collaborations between blood centers and transfusion services. *Transfusion*, 49(10):2040–2047, 2009.

[42] N. Ganikhodjaev, J.I. Daoud, and M. Usmanova. Linear and nonlinear models of heredity for blood groups and rhesus factor. *Journal of Applied Sciences*, 10(16):1748–1754, 2010.

[43] N. Ganikhodjaev, J.I. Daoud, and M. Usmanova. Stochastic models of heredity rhesus factor. *Australian Journal of Basic and Applied Sciences*, 4 (8):3306–3310, 2010.

[44] N. Ganikhodjaev, M. Saburov, and U. Jamilov. Mendelian and non-mendelian quadratic operators. *Applied Mathematics & Information Sciences*, 7(5):1721–1729, 2013.

[45] R. Ganikhodzhaev, F. Mukhamedov, and U. Rozikov. Quadratic stochastic operators and processes: results and open problems. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 14(2):279–335, 2011.

[46] E.A. Gehrie and C.A. Tormey. The influence of clinical and biological factors on transfusion-associated non-ABO antigen alloimmunization: responders, hyper-responders, and non-responders. *Transfusion Medicine and Hemotherapy*, 41(6):420–429, 2014.

[47] E.A. Gehrie, P.M. Ness, E.M. Bloch, S. Kacker, and A.A.R. Tobian. Medical and economic implications of strategies to prevent alloimmunization in sickle cell disease. *Transfusion*, 57(9), 2017.

[48] P.L.F. Giangrande. The history of blood transfusion. *British Journal of Haematology*, 110(4):758–767, 2000.

[49] P.W. Glynn. A GSMP formalism for discrete event systems. *Proceedings of the IEEE*, 77(1):14–23, 1989.

[50] S.K. Goyal and B.C. Giri. Recent trends in modeling of deteriorating inventory. *European Journal of Operational Research*, 134(1):1–16, 2001.

[51] S.C. Graves. Simple analytical models for perishable inventory systems. Technical Report No. 141, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, 1978.

[52] S. Gunpinar and G. Centeno. Stochastic integer programming models for reducing wastages and shortages of blood products at hospitals. *Computers & Operations Research*, 54:129–141, 2015.

[53] S.W. Guo and E.A. Thompson. Performing the exact test of hardy-weinberg proportion for multiple alleles. *Biometrics*, 48(2):361–372, 1992.

[54] R. Haijema, J. van der Wal, and N.M. van Dijk. Blood platelet production: Optimization by dynamic programming and simulation. *Computers & Operations Research*, 34(3):760–779, 2007.

[55] R. Haijema, N.M. van Dijk, J. van der Wal, and C.S. Sibinga. Blood platelet production with breaks: optimization by SDP and simulation. *International Journal of Production Economics*, 121(2):464–473, 2009.

[56] B. Hamdan and A. Diabat. A two-stage multi-echelon stochastic blood supply chain problem. *Computers & Operations Research*, 101:130–143, 2019.

[57] G. Hashmi, T. Shariff, Y. Zhang, J. Cristobal, C. Chau, M. Seul, P. Vissava-jjhala, C. Baldwin, K. Hue-Roye, D. Charles-Pierre, C. Lomas-Francis, and M.E. Reid. Determination of 24 minor red blood cell antigens for more than 2000 blood donors by high-throughput DNA analysis. *Transfusion*, 47(4): 736–747, 2007.

[58] J.E. Hendrickson and M. Delaney. Hemolytic disease of the fetus and new-born: modern practice and future investigations. *Transfusion Medicine Reviews*, 30(4):159–164, 2016.

[59] J.M. Higgins and S.R. Sloan. Stochastic modeling of human rbc alloim-munization: evidence for a distinct population of immunologic responders. *Blood*, 112(6):2546–2553, 2008.

[60] S.M. Hosseini-Motlagh, M.R.G. Samani, and S. Homaei. Blood supply chain management: robust optimization, disruption risk, and blood group com-patibility (a real-life case). *Journal of Ambient Intelligence and Humanized Computing*, pages 1–20, 2019.

[61] N. Jain. How precision medicine will change the future of healthcare, 2011. https://www.weforum.org/agenda/2019/01/why-precision-medicine-is-the-future-of-healthcare/ [Accessed: 17 January 2020].

[62] L. Janssen, T. Claus, and J. Sauer. Literature review of deteriorating in-ventory models by key topics from 2012 to 2015. *International Journal of Production Economics*, 182:86–112, 2016.

[63] M.P. Janssen, A.J.W. Tilborgh, K.M.K. Vooght, A.G. Bokhorst, and J.C. Wiersum-Osselton. Direct costs of transfusion reactions–an expert judge-ment approach. *Vox Sanguinis*, 113(2):143–151, 2017.

[64] H. Kaspi and D. Perry. Inventory systems of perishable commodities. *Ad-vances in Applied Probability*, 15(3):674–685, 1983.

[65] H.G. Klein, W.A. Flegel, and C. Natanson. Red blood cell transfusion: precision vs imprecision medicine. *JAMA*, 314(15):1557–1558, 2015.

[66] J.M. Koelewijn, M. de Haas, T.G.M. Vrijkotte, G.J. Bonsel, and C.E. van der Schoot. One single dose of 200 μg of antenatal RhIG halves the risk of anti-D immunization and hemolytic disease of the fetus and newborn in the next pregnancy. *Transfusion*, 48(8):1721–1729, 2008.

[67] G.F. Körmöczi and W.R. Mayr. Responder individuality in red blood cell alloimmunization. *Transfusion Medicine and Hemotherapy*, 41(6):446–451, 2014.

[68] K. Landsteiner. On agglutination of normal human blood [translation of: Über Agglutinationserscheinungen normalen menschlichen Blutes. *Wiener Klinische Wochenschrift*, 14:1132–1134, 1901.]. *Transfusion*, 1(1):5–8, 1961.

[69] W.J. Lane, C.M. Westhoff, N.S. Gleadall, M. Aguad, R. Smeland-Wagman, S. Vege, D.P. Simmons, H.H. Mah, M.S. Lebo, K. Walter, N. Soranzo, E. Di Angelantonio, J. Danesh, D.J. Roberts, N.A. Watkins, W.H. Ouwehand, A.S. Butterworth, R.M. Kaufman, H.L. Rehm, L.E. Silberstein, and R.C. Green. Automated typing of red blood cell and platelet antigens: a whole-genome sequencing study. *The Lancet Haematology*, 5(6):e241–e251, 2018.

[70] M. LaSalle-Williams, R. Nuss, T. Le, L. Cole, K. Hassell, J.R. Murphy, and D.R. Ambruso. Extended red blood cell antigen matching for transfusions in sickle cell disease: a review of a 14-year experience from a single center (cme). *Transfusion*, 51(8):1732–1739, 2011.

[71] P. Levine and R.E. Stetson. An unusual case of intra-group agglutination. *Journal of the American Medical Association*, 113(2):126–127, 1939.

[72] S.A. Lippman. Applying a new device in the optimization of exponential queuing systems. *Operations Research*, 23(4):687–710, 1975.

[73] S.M. Meinderts, J.W.R. Sins, K. Fijnvandraat, S.Q. Nagelkerke, J. Geissler, M.W. Tanck, C. Bruggeman, B.J. Biemond, A.W. Rijneveld, J.L.H. Kerkhoffs, S. Pakdaman, A. Habibi, R. van Bruggen, T.W. Kuijpers, F. Pirenne, and T.K. van den Berg. Nonclassical FCGR2C haplotype is associated with protection from red blood cell alloimmunization in sickle cell disease. *Blood*, 130(19):2121–2130, 2017.

[74] S. Meyer, N. Trost, B.M. Frey, et al. Parallel donor genotyping for 46 selected blood group and 4 human platelet antigens using high-throughput MALDI-TOF mass spectrometry. In *Molecular Typing of Blood Cell Antigens*, pages 51–70. Springer, 2015.

[75] A. Montpetit, M.S. Phillips, I. Mongrain, R. Lemieux, and M. St-Louis. High-throughput molecular profiling of blood donors for minor red blood cell and platelet antigens. *Transfusion*, 46(5):841–848, 2006.

[76] J.A. Morton and M.M. Pickles. Use of trypsin in the detection of incomplete anti-Rh antibodies. *Nature*, 159:779, 1947.

[77] A. Nagurney, A.H. Masoumi, and M. Yu. Supply chain network operations management of a blood banking system with cost and risk minimization. *Computational Management Science*, 9(2):205–231, 2012.

[78] S. Nahmias. Perishable inventory theory: A review. *Operations Research*, 30(4):680–708, 1982.

[79] S. Nahmias. *Perishable inventory systems*. Springer Science & Business Media, New York, NY, 2011.

[80] M.F. Neuts. *Matrix-geometric solutions in stochastic models: an algorithmic approach*. Johns Hopkins University Press, Baltimore, MD, 1994.

[81] R.S. Nickel, J.E. Hendrickson, R.M. Fasano, E.K. Meyer, A.M. Winkler, M.M. Yee, P.A. Lane, Y.A. Jones, F.D. Pashankar, T. New, C.D. Josephson, and Stowell S.R. Impact of red blood cell alloimmunization on sickle cell disease mortality: a case series. *Transfusion*, 56(1):107–114, 2016.

[82] A.R. Odoni. On finding the maximal gain for markov decision processes. *Operations Research*, 17(5):857–860, 1969.

[83] A.F. Osorio, S.C. Brailsford, and H.K. Smith. A structured review of quantitative models in the blood supply chain: a taxonomic framework for decision-making. *International Journal of Production Research*, 53(24):7191–7212, 2015.

[84] M. Parlar, D. Perry, and W. Stadje. FIFO versus LIFO issuing policies for stochastic perishable inventory systems. *Methodology and Computing in Applied Probability*, 13(2):405–417, 2011.

[85] W.P. Pierskalla. Supply chain management of blood banks. In *Operations research and health care: a handbook of methods and applications*, pages 103–145. Springer Science & Business Media, New York, NY, 2004.

[86] A. Pirabán, W.J. Guerrero, and N. Labadie. Survey on blood supply chain management: models and methods. *Computers & Operations Research*, 112, 2019.

[87] W.B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*. John Wiley & Sons, Hoboken, NJ, 2007.

[88] G.P. Prastacos. Blood inventory management: an overview of theory and practice. *Management Science*, 30(7):777–800, 1984.

[89] M.L. Puterman. *Markov Decision Processes.: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Hoboken, NJ, 2014.

[90] F. Raafat. Survey of literature on continuously deteriorating inventory models. *Journal of the Operational Research Society*, 42(1):27–37, 1991.

[91] E. Rachelson, G. Quesnel, F. Garcia, and P. Fabiani. A simulation-based approach for solving generalized semi-markov decision processes. In *European Conference on Artificial Intelligence*, pages 583–587, 2008.

[92] S. Rajendran and A.R. Ravindran. Platelet ordering policies at hospitals using stochastic integer programming model and heuristic approaches to reduce wastage. *Computers & Industrial Engineering*, 110:151–164, 2017.

[93] G. Ramsey and S.J. Smietana. Long-term follow-up testing of red cell alloantibodies. *Transfusion*, 34(2):122–124, 1994.

[94] M.E. Reid, C. Lomas-Francis, and M.L. Olsson. *The blood group antigen factsbook*. Academic Press, London, UK, 2012.

[95] R. Reverberi. The persistence of red cell alloantibodies. *Blood Transfusion*, 6(4):225–234, 2008.

[96] M.A.M. Rogers, J.M. Rohde, and N. Blumberg. Haemovigilance of reactions associated with red blood cell transfusion: comparison across 17 countries. *Vox Sanguinis*, 110(3):266–277, 2016.

[97] S.M. Ross. *Introduction to probability models*. Academic press, Burlington, MA, 2014.

[98] T. Sadykov. Polynomial dynamics of human blood genotypes frequencies. *Journal of Symbolic Computation*, 79:342–355, 2017.

[99] Sanquin. Annual report 2018, 2018. https://www.sanquin.nl/en/about-sanquin/annual-reports [Accessed: 17 January 2020].

[100] B. Santos, R. Portugal, C. Nogueira, and M. Loureiro. Hyperhemolysis syndrome in patients with sickle cell anemia: report of three cases. *Transfusion*, 55(6pt2):1394–1398, 2015.

[101] V. Sarhangian, H. Abouee-Mehrizi, O. Baron, and O. Berman. Threshold-based allocation policies for inventory management of red blood cells. *Manufacturing & Service Operations Management*, 20(2):347–362, 2017.

[102] H. Schonewille, H.L. Haak, and A.M. van Zijl. RBC antibody persistence. *Transfusion*, 40(9):1127–1131, 2000.

[103] H. Schonewille, L.M.G. van de Watering, and A. Brand. Additional red blood cell alloantibodies after blood transfusions in a nonhematologic alloimmunized patient cohort: is it time to take precautionary measures? *Transfusion*, 46(4):630–635, 2006.

[104] H. Schonewille, Á. Honohan, L.M.G. van de Watering, F. Hudig, P.A. te Boekhorst, A.W.M.M. Koopman-van Gemert, and A. Brand. Incidence of alloantibody formation after ABO-D or extended matched red blood cell transfusions: a randomized trial (MATCH study). *Transfusion*, 56(2):311–320, 2016.

[105] R.F. Serfozo. An equivalence between continuous and discrete time markov decision processes. *Operations Research*, 27(3):616–620, 1979.

[106] S.H.W. Stanger, R. Wilding, N. Yates, and S. Cotton. What drives perishable inventory management performance? Lessons learnt from the UK blood supply chain. *Supply Chain Management*, 17(2):107–123, 2012.

[107] S.H.W. Stanger, N. Yates, R. Wilding, and S. Cotton. Blood inventory management: hospital best practice. *Transfusion Medicine Reviews*, 26(2):153–163, 2012.

[108] J.R. Storry, F.B. Clausen, L. Castilho, Q. Chen, G. Daniels, G. Denomme, W.A. Flegel, C. Gassner, M. de Haas, C. Hyland, et al. International society of blood transfusion working party on red cell immunogenetics and blood group terminology: report of the dubai, copenhagen and toronto meetings. *Vox Sanguinis*, 114(1):95–102, 2019.

[109] Z. Tatari-Calderone, H. Gordish-Dressman, R. Fasano, M. Riggs, C. Fortier, A.D. Campbell, D. Charron, V.R. Gordeuk, N.L.C. Luban, S. Vukmanovic, and R. Tamouza. Protective effect of HLA-DQB1 alleles against alloimmunization in patients with sickle cell disease. *Human Immunology*, 77(1):35–40, 2016.

[110] H.C. Tijms. *A first course in stochastic models*. John Wiley & Sons Ltd, Chichester, UK, 2003.

[111] H.C. Tijms. One-step improvement ideas and computation aspects. In R.J. Boucherie and N.M. Van Dijk, editors, *Markov decision processes in practice*. Springer, Cham, Switzerland, 2017.

[112] H. Tinegate, K. Pendry, M. Murphy, et al. Where do all the red blood cells (RBCs) go? Results of a survey of RBC use in England and North Wales in 2014. *Transfusion*, 56(1):139–145, 2016.

[113] C.A. Tormey and J.E. Hendrickson. Transfusion-related red blood cell alloantibodies: induction and consequences. *Blood*, 133(17):1821–1830, 2019.

[114] L.N. Trefethen and D. Bau III. *Numerical linear algebra*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997.

[115] N.M. Van Dijk. *Controlled Markov processes: time-discretization*. CWI (Centrum voor Wiskunde en Informatica) Tract No. 11, 1984.

[116] N.M. Van Dijk. *Queueing networks and product forms: a systems approach*, volume 4. John Wiley & Sons Ltd, Chichester, UK, 1993.

[117] N.M. Van Dijk, R. Haijema, J. Van Der Wal, and C.S. Sibinga. Blood platelet production: a novel approach for practical optimization. *Transfusion*, 49(3): 411–420, 2009.

[118] N.M. van Dijk, S.P.J. van Brummelen, and R.J. Boucherie. Uniformization: Basics, extensions and applications. *Performance evaluation*, 118: 8–32, 2018.

[119] A. van Dongen, M. Mews, W.L.A.M. de Kort, and E. Wagenmans. Missing minorities: A survey based description of the current state of minority blood donor recruitment across 23 countries. *Diversity & Equality in Health and Care*, 13:138–145, 2016.

[120] A.J. van Gammeren, A.G. van den Bos, N. Som, C. Veldhoven, R.C.R.M. Vossen, and C.C. Folman. A national Transfusion Register of Irregular Antibodies and Cross (X)-match Problems: TRIX, a 10-year analysis. *Transfusion*, 59(8):2559–2566, 2019.

[121] L.R. van Hoeven, B.H. Hooftman, M.P. Janssen, M.C. de Bruijne, K.M.K. de Vooght, P. Kemper, and M.M.W. Koopman. Protocol for a national blood transfusion data warehouse from donor to recipient. *BMJ Open*, 6(8): e010962, 2016.

[122] B. Veldhuisen, C.E. van der Schoot, and M. de Haas. Blood group genotyping: from patient to high-throughput donor screening. *Vox Sanguinis*, 97 (3):198–206, 2009.

[123] E.P. Verduin, A. Brand, and H. Schonewille. Is female sex a risk factor for red blood cell alloimmunization after transfusion? A systematic review. *Transfusion Medicine Reviews*, 26(4):342–353, 2012.

[124] E.P. Verduin, A. Brand, R.A. Middelburg, and H. Schonewille. Female sex of older patients is an independent risk factor for red blood cell alloimmunization after transfusion. *Transfusion*, 55(6pt2):1478–1485, 2015.

[125] N. Vreeswijk, J. Jongerius, A. Van Weert, and H. Bos. Vijftien miljoen typeringen van het donorbestand biedt tijdwinst en ondersteunt adequate preventie van immunisatie rondom bloedtransfusie. *NederlandsTijdschrift voor Klinische Chemie en Laboratoriumgeneeskunde*, 40:64–67, 2015.

[126] C.M. Westhoff. Blood group genotyping. *Blood*, 133(17):1814–1820, 2019.

[127] L.M. Williamson and D.V. Devine. Challenges in the management of the blood supply. *The Lancet*, 381(9880):1866–1875, 2013.

[128] W.L. Winston. *Operations research: applications and algorithms*, volume 3. Brooks/Cole - Thomson Learning, Belmont, CA, 2004.

[129] World Health Organization. Blood safety and availability., 2019. `https://www.who.int/news-room/fact-sheets/detail/blood-safety-and-availability` [Accessed: 17 January 2020].

[130] K. Yazdanbakhsh. Mechanisms of sickle cell alloimmunization. *Transfusion Clinique et Biologique*, 22(3):178–181, 2015.

[131] H.L.S. Younes and R.G. Simmons. Solving generalized semi-markov decision processes using continuous phase-type distributions. *Proceedings of the AAAI*, 4:742–748, 2004.

[132] B. Zahiri and M.S. Pishvaee. Blood supply chain network design considering blood group compatibility under uncertainty. *International Journal of Production Research*, 55(7):2013–2033, 2016.

[133] S. Zalpuri, J.J. Zwaginga, and J.G. van der Bom. Risk Factors for alloimmunisation after red blood Cell Transfusions (R-FACT): a case cohort study. *BMJ Open*, 2(3):e001150, 2012.

[134] S. Zalpuri, D. Evers, J.J. Zwaginga, H. Schonewille, K.M.K. Vooght, S. Cessie, and J.G. Bom. Immunosuppressants and alloimmunization against red blood cell transfusions. *Transfusion*, 54(8):1981–1987, 2014.

[135] C. Zwiers, J.M. Koelewijn, L. Vermij, J.H.J. van Sambeeck, D. Oepkes, M. de Haas, and C.E. van der Schoot. ABO incompatibility and RhIG immunoprophylaxis protect against non-D alloimmunization by pregnancy. *Transfusion*, 58(7):1611–1617, 2018.

Summary

Introduction

A blood transfusion is a safe, common, and potentially life-saving medical procedure in which one or multiple blood components, originating from a donor, are inserted into the bloodstream of a transfusion recipient. The recipient might, for example, suffer from a haematological disorder (leukaemia, sickle cell disease, thalassemia) or surgical procedure (organ transplant, heart surgery). With respect to transfusion of red blood cells, it is important that the blood groups of the donor and transfusion recipient match.

The blood group of an individual is determined by the presence or absence of antigens on the surface of the red blood cells. If a particular antigen is present on red blood cells of a donor but absent on the red blood cells of the transfusion recipient, the immune system of the transfusion recipient may produce against this foreign antigen (alloimmunization) These antibodies will cause problems during a subsequent transfusion or, in the case of a female recipient, a future pregnancy. A response of the immune system could have been prevented by selecting red blood cells lacking the relevant antigen.

Donor selection

As blood groups are inherited, the recruitment of donor relatives, for example siblings, is expected to be effective, since the probability of a similar rare blood group is likely. However, the likelihood differs between blood groups and is not commonly available. This paper provides a unified mathematical formulation to calculate such likelihoods. From a mathematical and probabilistic point of view, it is shown that these likelihoods can be obtained from the computation of a stationary genotype distribution. This, in turn, can be brought down to a system of quadratic stochastic operators. A generic mathematical approach is presented which directly leads to a stationary genotype distribution for arbitrary blood groups.

Inventory allocation

The ABO blood group classification as well as the Rhesus-D factor are well known, which includes three blood group antigens (A, B, D). Within the field of transfusion medicine, however, more than 15 different blood group antigens are distinguished. Ideally, each transfusion recipient is matched for these clinically relevant blood group antigens. If there were unlimited quantities of red blood cell units in inventory, this would in principle be no problem. In practice, however, inventory size are finite.

We have developed a new mathematical framework that can be applied for an arbitrary set of antigens. It determines which red blood cell units should be issued from inventory, such that 1) requests from hospitals can be satisfied with antigen compatible red blood cell units 2) shortages for future requests are avoided, and 3) outdating is prevented. The potential practical performance of the optimization model is evaluated by simulation and is based on historical data on supply and demand of red blood cell units in the Netherlands. The results show that most requests can be satisfied with antigen identical red blood cell units, while shortages and outdating can be kept restricted.

In practice, shortages are solved by reducing the number of antigens that are matched. However, the probablity of alloimmunization increases. This probability can be minimized by determining an optimal sequence for antigen exclusion in case insufficient matching red blood cell units are available from inventory.

Conclusion

At present, alloimmunization preventing matching strategies are only applied to specific transfusion recipient groups and include a limited number of red blood cell antigens. The general assumption is that providing extensively matched red blood cell units to all transfusion recipients is not feasible in practice. In this thesis, we investigate whether this assumption is valid and compute the proportion of alloimmunization that can be prevented when all donors and transfusion recipients are comprehensively typed.

Samenvatting

Introductie

Een bloedtransfusie is een veilige, veel voorkomende en soms levensreddende medische behandeling, waarbij rode bloedcellen afkomstig van een donor worden ingebracht in de bloedbaan van een transfusieontvanger die deze nodig heeft vanwege een chirurgische ingreep (orgaantransplantatie, hartchirurgie) of een hematologische aandoening (leukemie, sikkelcelanemie, thalassemie). Bij een bloedtransfusie is het van belang dat de bloedgroepen van de donor en transfusieontvanger zoveel mogelijk overeenkomen, of anders gezegd, dat hun bloedgroepen matchen.

De bloedgroep van een individu wordt bepaald door de aan –of afwezigheid van antigenen op de celwand van de rode bloedcellen. Indien een bepaald antigeen aanwezig is op de rode bloedcel van een donor, maar afwezig op de rode bloedcel van de transfusieontvanger, dan kan het immuunsysteem van de transfusie ontvanger overgaan tot het aanmaken van antistoffen tegen dit lichaamsvreemde antigeen (alloimmunisatie). Deze antistoffen kunnen vervolgens problemen veroorzaken bij een volgende bloedtransfusie, of in het geval van een vrouwelijke ontvanger, een toekomstige zwangerschap. Een reactie van het immuunsysteem kan echter voorkomen worden door uit de beschikbare voorraad –indien mogelijk– rode bloedcellen te selecteren waarbij het desbetreffende antigeen ook ontbreekt.

Donor selectie

Huidige donor wervingsstrategieën zijn gebaseerd op historische matchingsstrategieën en kunnen niet voldoen aan de vraag naar uitgebreid gematchte rode bloedceleenheden. Omdat bloedgroepen erfelijk zijn, is de kans op het vinden van een vergelijkbaar antigeenprofiel bij naaste familieleden veel groter dan bij een willekeurig persoon. Voor selectieve werving van potentiële donors is het interessant te weten hoe groot deze kans als functie van de familierelatie is. Dit lijkt relatief eenvoudig, maar de a priori kansen zijn echter onbekend en zijn bepaald door het oplossen van een stelsel van kwadratische vergelijkingen.

Voorraad management

Over het algemeen is de ABO-bloedgroep classificatie evenals de Rhesus-D factor welbekend, welke een drietal antigenen betreft (A, B, D). Binnen de transfusiegeneeskunde onderscheidt met echter meer dan 15 verschillende, klinisch relevante, bloedgroepantigenen. Idealiter zou men bij een bloedtransfusie met al deze antigenen rekening willen houden. Indien er onbeperkte hoeveelheden rode bloedceleenheden voorradig zouden zijn, is dat in principe geen probleem. In de praktijk zijn voorraden echter eindig.

We hebben een nieuw wiskundig optimalisatie model ontwikkeld dat kan worden toegepast voor een willekeurige verzameling antigenen. Het bepaalt welke rode bloedceleenheden uit de voorraad geselecteerd moeten worden zodat 1) aanvragen van ziekhuizen kunnen worden voldaan met antigeen compatibele eenheden, 2) tekorten worden vermeden en 3) zo min mogelijk eenheden over de houdbaarheidsdatum gaan. De werking van het wiskundig optimalisatie model wordt geëvalueerd door middel van simulaties en is gebaseerd op historisch data omtrent vraag en aanbod van rode bloedceleenheden in Nederland. Uit de resultaten blijkt dat een groot aantal van de aanvragen van ziekenhuizen kunnen worden voldaan met antigeen identieke rode bloedceleenheden, en dat de kans op tekorten en overschotten hierbij klein is.

In de praktijk worden tekorten opgelost door het aantal antigenen waarop wordt gematcht te reduceren. Echter neemt de kans op de vorming van alloantistoffen hierdoor toe. Deze kans kan geminimaliseerd worden door een optimale volgorde te bepalen waarop antigenen worden weggestreep uit de matchingsstrategie.

Conclusie

Momenteel worden strategieën ter voorkoming van antistofvorming alleen toegepast op specifieke groepen transfusieontvangers en omvatten slechts een beperkt aantal rode bloedcelantigenen. De algemene aanname is dat het verstrekken van uitgebreid gematchte rode bloedceleenheden aan alle transfusieontvangers praktisch niet haalbaar is. In dit proefschrift onderzoeken we of deze aanname klopt en berekenen we het percentage antistoffen dat voorkomen kan worden wanneer alle donoren en transfusieontvangers volledig zijn getypeerd.

Dankwoord

Mijn PhD-traject doet me denken aan een andere uitdaging die ik in de afgelopen jaren ben aangegaan, namelijk het deelnemen aan obstacle runs en het lopen van halve marathons. Hoewel het misschien lijkt alsof je het individueel doet, kan je het zeker niet alleen. Hiervoor wil ik een aantal mensen in het bijzonder bedanken.

Promotoren

Ik wil graag beginnen met mijn promotoren. Mart, Nico, Wim, jullie hebben alle drie verschillende persoonlijkheden en achtergronden. Gecombineerd is het een mooie formule voor een geweldig team van promotoren! Bedankt voor de begeleiding in de afgelopen jaren.

Mart bedankt voor je optimisme, oprechtheid en motivatie. Je hebt me geïntroduceerd binnen Sanquin en geholpen om mijn onderzoek echt relevant te maken. Dit heeft mij enorm veel energie gegeven om continu het beste uit mezelf te halen. Je wist mij te overtuigen dat mijn schrijf –en presenteervaardigheden hartstikke goed waren en hierdoor mijn onzekerheid weg te nemen (of in ieder geval te verminderen). Daarnaast heb ik het gevoel dat ik op je kon terugvallen als ik er zelf niet uit kwam, bedankt hiervoor.

Nico bedankt voor jouw betrokkenheid, flexibiliteit en zorgzaamheid. Je nam tijd om uren lang met mij te overleggen, samen te werken en jouw expertise, stochastische processen, met mij te delen. Ik heb met enorm veel plezier van je geleerd. Vaak waren onze overleggen bij Sanquin, maar af en toe ook in de trein van Hengelo naar het westen van het land. We hebben de meeste artikelen afgerond in de trein. Daarbij was het eindstation niet afhankelijk van de locatie waar we heen wilden, maar van hoeveel er nog moest gebeuren voordat we klaar waren. Ook hebben we vaak over mijn sportieve ambities gesproken, je was altijd geïnteresseerd. Ik kijk ernaar uit om nog een keer samen in Limburg te gaan fietsen.

Wim bedankt voor je enthousiasme en interesse. Ik ben enorm onder de indruk van jouw wiskundige kennis. Jij wist de formules altijd goed te doorgronden en precies de juiste vragen te stellen, zodat ik het nog duidelijker op kon schrijven.

Commissie leden

Dear committee members, Prof. dr. J. Blake, Prof. dr. R. Boucherie, Prof. dr. M. de Haas, Prof. dr. Ir. E. Hans, Prof. dr. ir. D. den Hertog, and Prof. dr. E. van der Schoot, thank you for joining the doctoral committee, and for investing your time and effort to read, assess, and discuss my dissertation.

Sanquin collega's

Ik wil Sanquin en het BloodMatch-team bedanken voor het creëren van de onderzoeksrichting waaraan ik in de afgelopen jaren heb mogen werken. De vele discussies die we aan het begin van mijn PhD-traject hebben gevoerd, hebben ervoor gezorgd dat ik, als wiskundige, me binnen no-time thuis voelde in de wereld van de transfusiegeneeskunde. Hierbij wil ik een aantal mensen in het bijzonder bedanken.

Henk, bedankt voor je kritische blik en het veelvuldig nalezen van mijn artikelen. Ik vind het prachtig om te zien hoe jij met het schrappen van een paar woorden, mijn zinnen wetenschappelijk kon laten klinken. Je had altijd tijd voor me om jouw commentaar met mij te bespreken. Ellen, ondanks je drukke agenda wisten we altijd (vaak aan het einde van de dag als iedereen al naar huis was) een moment te vinden om mijn onderzoek te bespreken. Deze gesprekken hebben ervoor gezorgd dat de praktische relevantie van mijn onderzoek nog duidelijker voor mij is geworden. Jouw enthousiasme motiveerde me enorm. Masja, bedankt dat jij me liet inzien dat het ultieme doel van de bloedbank niet is om aan alle vraag van ziekenhuizen te voldoen, maar om de eenheden zo over ziekenhuizen te verdelen dat elke transfusieontvanger passend bloed krijgt. Marian, we reisde weleens samen met de metro naar Amsterdam Zuid en hadden dan gesprekken over hoe wiskunde de bloedbank zou kunnen helpen bij het optimaliseren van haar logistieke processen. Bedankt voor jouw open mindset en de interessante gesprekken die we hierover hebben gehad.

TTA collega's, Mart, Anneke, Loan, Syeldy, Roxy, Shannon, Marieke, we waren een mooie internationale groep, waarbij ik de persoonlijk gesprekken erg heb gewaardeerd. Daarnaast heb ik genoten van de verschillende internationale, soms veganistische, diners. Het was altijd heerlijk om de verschillende werelden te proeven.

Donorstudies collega's, Wim, Anton, Katja, Eva, Lisanne, Femmeke, Saurabh, Marloes, Steven, Bas, Veerle, Maike, Franke, Mariëlle, Marie, Jan, Winesh, bedankt dat ik ook onderdeel uit heb mogen maken van deze groep. De lunches, uitjes, diners en (niet te vergeten) borrels waren altijd erg gezellig.

Maurits, Sem, Rosa, Tjeerd, Lisa, Tiffany, ik waardeer de vriendschap die we in de afgelopen jaren hebben opgebouwd. De traditionele pannenkoeken avonden

waren erg gezellig. Over de taakverdeling hoefde we niet na te denken, deze stond namelijk al vast. Ook de OIO-uitjes en later klimavonden waren een succes. Maurits, Sem, jullie hebben ons als eerste verlaten, maar een belangrijke basis gelegd voor onze OIO-groep, bedankt hiervoor. Rosa, Tjeerd, ik zie jullie (samen met Katja) als mijn Sanquin hardloopmaatjes. Het zou leuk zijn als we binnenkort weer een keer samen kunnen deelnemen aan een hardloopwedstrijd. Lisa, Tiffanny, bedankt voor fijne koffie momenten in de vroege ochtend en voor de gezellige wandelingen naar de metro aan het einde van de dag.

UT Collega's Richard, Erwin, ik wil jullie bedanken voor het mogelijk maken van mijn promotieonderzoek en voor de prettige werksfeer die jullie binnen CHOIR hebben weten te creëren. Het was fijn om onderdeel uit te maken van deze onderzoeksgroep. Ook al was ik niet werkzaam in een ziekenhuis, ik heb enorm veel geleerd over de complexe logistieke processen die hier plaatsvinden. Tegen al mijn CHOIR collega's Nico, Derya, Aleida, Maartje, Gréanne, Sem, Nardo, Ingeborg, Bruno, Thomas, Maarten, Jasper, Eline, Marelise, Jasmijn, Nikki en alle andere MOR collega's wil graag zeggen dat ik elke week met veel plezier naar Enschede afreisde. Ik heb erg genoten van de congressen, barbecues, lunchwandelingen, uitjes, en wiskundige discussies bij de koffieautomaat.

Corine, Xinwei, en later Anne, jullie hebben er als kamergenoten voor gezorgd dat ik me thuis voelde op mijn werkplek op de UT. Corine, ook al zei ik dat ik niet voor de gezelligheid kwam, wil ik je toch bedanken voor de vele leuke gesprekken die wij met elkaar hebben gevoerd, eerst op de UT en later ook in de trein terug naar huis. Xinwei, bedankt voor zowel de rust als gezelligheid die jij in ons kantoor bracht. Ik vond het altijd erg leuk om Nederlands met je de praten. Anne, bedankt dat ik je altijd mocht bellen als ik weer eens met een prangende vraag zat hoe ik iets het beste voor de promotieceremonie kon regelen.

Familie en vrienden

Mijn vrienden van de middelbare school en universiteit bedankt voor jullie interesse in mijn promotie, leuke afleiding in de weekenden en begrip als ik traag reageerde op jullie berichten. Jan en Petra, bedankt voor jullie gastvrijheid en voor het lekkere stukje Limburgse vlaai dat elk weekend op het menu stond. Tante Maria, bedankt voor het prachtige ontwerp van de kaft van mijn proefschrift.

Papa en mama, sinds de middelbare school hebben jullie me al vrijgelaten om mijn eigen weg te kiezen. Bedankt dat jullie altijd achter de keuzes stonden die ik heb gemaakt en dat de deur in Reusel altijd voor mij open staat. Coen, Frits, fijn dat jullie mij als paranimf willen bijstaan bij de verdediging. Ellen leuk om te zien dat jij ook je weg naar het onderzoek hebt gevonden. Het is bijzonder, zeker in deze tijd, dat jullie allemaal bij mijn promotie aanwezig kunnen zijn. Dit vind ik ontzettend fijn!

Lieve Chris, het samenwonen met een PhD student is niet altijd even makkelijk. Op een gegeven moment hebben we de vraag "hoe gaat het met je PhD", die jij elke avond stelde, ingewisseld voor de vraag "hoe gaat het met je collega's". Dit leverde na een lange dag hard werken veel leukere gesprekken op. Bedankt dat je me al die tijd hebt gesteund.

Joost van Sambeeck
Eindhoven, November 2020

About the author

Joost van Sambeeck was born on September 28, 1991 in Reusel, the Netherlands. After completing primary school at the Leilinde in Reusel, he went to Pius-X College in Bladel, where he finished his preparatory university education (in Dutch: Gymnasium) in 2009, in Science and Health. In 2013, he completed the Bachelor's programme in Econometrics and Operations Research at Tilburg University. During the Bachelor's programme, he developed a passion for mathematical optimization and therefore decided to pursue a Master's in Operations Research and Management Science, also at Tilburg university, which he completed in 2014.

Following a recommendation from Prof. dr. ir. Dick den Hertog – his former supervisor during his Master's thesis on the robust facility location problem – Joost applied for a Ph.D. position at the department of Transfusion Technology Assessment at Sanquin Research, and the Centre of Healthcare Operations Improvement and Research (CHOIR) at the University of Twente. He started at the end of 2014 under supervision of Dr. Mart Janssen, Prof. dr. Nico van Dijk, and Prof. dr. Wim de Kort. During his research he was immersed into the world of blood group matching and showed that extended red blood cell matching is feasible in practice. His Ph.D. research resulted in five scientific papers, two non-scientific papers and two scientific awards (i.e., Best Poster Award at NVB-TRIP symposium 2016 and Best Poster Award at Sanquin Science day 2017) and culminates with this thesis and the defence on December 9th, 2020.

He continues his career and passion for mathematical optimization as a consultant at CQM in Eindhoven, where he started in April this year.

After being born, raised and having lived in the province of Noord-Brabant for more than 20 years, Joost moved for his PhD to the middle of the Netherlands and lived in Utrecht and Gouda. He travelled every day by public transport to Amsterdam or Enschede. Recently, he moved back to Noord-Brabant and now lives in Eindhoven, 15 minutes walking from the office.

List of publications

Van Sambeeck, J.H.J., Janssen, M.P., Ligthart, P., de Kort, W.L.A.M., & Van Dijk, N.M. (2016). Heb jij dat ook van je broer of zus? *STAtOR, 17*(2), 26-31.

Van Sambeeck, J.H.J., Janssen, M.P., Schonewille, H., de Kort, W.L.A.M., & Van Dijk, N.M. (2018). Blood group probabilities by next of kin. *Probability in the Engineering of Informational Sciences*, 1-21.
(Basis for Chapter 3)

Van Sambeeck, J.H.J., De Wit, P.D., Luken, J., Veldhuisen, B., Van den Hurk, K., Van Dongen, A., Koopman, M.M.W., Van Kraaij, M.G.J., Van der Schoot, C.E., Schonewille, H., De Kort, W.L.A.M., & Janssen, M.P., (2018). A conceptual framework for optimizing blood matching strategies: balancing patient complications against total costs incurred. *Frontiers in Medicine, 5*(199).
(Basis for Chapter 2)

Zwiers, C., Koelewijn, J.M., Vermij, L., **Van Sambeeck, J.H.J.**, Oepkes, D., De Haas, M., & Van der Schoot, C.E. (2018). ABO incompatibility and RhIG immunprophylaxis protect against non-D alloimmunization by pregnancy. *Transfusion, 58*(7) 1611-1617.

Van Sambeeck, J.H.J., Van der Schoot, C.E., Van Dijk, N.M., Schonewille, H., & Janssen, M.P. (2019). Mastermind voor BloodMatching. *STAtOR, 20*(2), 26-29.

Van Sambeeck, J.H.J., Van Brummelen, S.P.J., Van Dijk, N.M., & Janssen, M.P. (2020). Blood group specific issuing policies to improve inventory management of red blood cells. *Submitted.*
(Basis for Chapters 4 and 6)

Van Sambeeck, J.H.J., Van der Schoot, C.E., Van Dijk, N.M., Schonewille, H., & Janssen, M.P. (2020). Mathematical optimization for alloimmunization prevention. *Submitted.*
(Basis for Chapter 7)

Van Sambeeck J.H.J., Van Dijk, N.M, (2020). A microscopic mathematical description for optimal blood issuing. *Working Paper.*
(Basis for Chapter 5)

OPTIMAL RED BLOOD CELL MATCHING

Over the past years, extended antigen matching has become standard practice for some transfusion recipients and has proven to reduce alloimmunization risks. Inexpensive molecular typing has the potential to provide extended antigen matched red blood cell units to all transfusion recipients. It will alleviate unnecessary suffering which can result from the currently applied one-size-fits-all approach and will reduce costs of treatment by eliminating unintended side-effects.

By applying Operations Research techniques, we show that inventory issuing policies can be improved and evaluate the practical feasibility of extended antigen matching. The developed mathematical models are applied to case studies executed at the Dutch blood bank Sanquin.

Sanquin   UNIVERSITY OF TWENTE.   CHOIR