Analysis and optimization of blood testing procedures

EURANDOM PREPRINT SERIES

2015-016

August, 2015

Analysis and optimization of blood testing procedures

S. Bar-Lev, O. Boxma, D. Perry, P. Vastazos
ISSN 1389-2355

ANALYSIS AND OPTIMIZATION OF BLOOD TESTING PROCEDURES

SHAUL K. BAR-LEV[1], ONNO BOXMA[2], DAVID PERRY[3], AND PADELIS VASTAZOS[4]

Abstract. This paper is devoted to the performance analysis and optimization of blood testing procedures. We present a queueing model of two queues in series, representing the two stages of a blood testing procedure. Service (testing) in stage 1 is performed in batches, whereas it is done individually in stage 2. Since particular elements of blood can only be stored and used within a finite time window, the sojourn time of blood units in the system of two queues in series is an important performance measure, which we study in detail. We also introduce a profit objective function, taking into account blood acquisition and screening costs as well as profits for blood units which were found uncontaminated and were tested fast enough. We optimize that profit objective function w.r.t. the batch size and the length of the time window.

1. Introduction

This paper is devoted to the performance analysis and optimization of blood testing procedures. We evaluate the performance of a few blood testing alternatives, mainly focussing on the aspect of time. That focus is instigated by the fact that certain blood components have a limited "shelf life", after which they can no longer safely be used.

One of the major issues in securing blood supply to patients worldwide is to provide blood of the best achievable quality, in the needed quantities. Blood is collected from human donors, by Blood Services worldwide, using specially designated blood bags and equipment, as "whole blood" units. In most developed countries these units are separated into different components: packed red blood cells, plasma, cryoprecipitate and platelets. The various components are subsequently stored according to the different blood types, under the corresponding temperatures, storage conditions and expiration dates. Each individual blood donation undergoes various tests, to define the donor's blood type and to detect agents that may cause transfusion-transmitted diseases.

Hospitals and other Transfusion centers order certain quantities of the different blood components from the Central/Regional Blood Services (CBS) according to their operational needs. The CBS each day has to decide which quantities of blood components of each type are sent to each hospital, based on their requests and on the need to keep a sufficient inventory.

[1]Department of Statistics, Haifa University, Haifa, Israel (barlev@stat.haifa.ac.il).

[2]Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands (o.j.boxma@tue.nl).

[3]Department of Statistics, Haifa University, Haifa, Israel (dperry@stat.haifa.ac.il).

[4]Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands (l.p.vastazos@hotmail.com).

Blood donations are performed in fixed donor sites or using blood mobiles. The bags containing the donated blood are transferred to the CBS laboratories for testing, processing, storage and supply.

In the CBS laboratories, each blood donation goes through multiple tests for the presence of various pathogens which are able to cause transfusion-transmitted diseases. In most countries it is mandatory to screen all blood donations for hepatitis B (HBV), hepatitis C (HCV), human immunodeficiency virus (HIV) and syphilis. One remarkable fact is that about 35 years ago, before HBV and HCV testing of blood donations was mandatory, about 20% of the hepatitis cases were caused by blood transfusions. In the US, the FDA has progressively strengthened the overlapping safeguards that protect patients from unsuitable blood and blood products. Blood donors are asked specific questions about risk factors that could affect the safety of the donation and are deferred from donation if risk factors are acknowledged. FDA also requires blood centers to maintain lists of unsuitable donors to prevent further donations from these individuals. After donation, the blood is tested for several infectious agents. All tests must be negative before the blood is suitable for transfusion. In addition to these safeguards, FDA has significantly increased its oversight of the blood industry. The agency inspects all blood facilities at least every two years, and 'problem' facilities are inspected more often. Blood establishments are now held to quality standards comparable to those expected of pharmaceutical manufacturers. The cost of this screening is rising in developed countries and is a major economic burden in developing countries. See Schottstedt et al. [13], Chiavetta et al. [4], Jackson et al. [11], Stramer et al. [14], Marshall et al. [12], Hourfar et al. [10], Ghandforoush and Sen [7] and Stramer et al. [15].

The blood laboratories have developed two different test procedures. The older one is called ELISA (Enzyme Linked Immuno-Sorbent Assay). This procedure detects virus-specific antibodies in the blood. The benefit of this procedure is that it has high sensitivity and specificity, so the blood sample could be screened properly. There is one disadvantage though. There are viruses such as HIV, for which the immune system requires a lot of time to develop a high concentration of antibodies. As a result, the ELISA test can not detect the virus during the first days (or weeks) after infection. Thus, the ELISA procedure has a lower analytic detection limit. Actually the period elapsing from the time a person is infected by some virus until antibodies can be detected, is called *window period*. Examples of average window periods for some viruses are: 22 days for HIV, 60 days for HBV and 70 days for HCV. During the window period, the ELISA method might provide wrong information about the blood sample. This problem was the motivation for the development of a new test procedure which is called PCR (Polymerase Chain Reaction). This test detects viral genetic material in the blood which is a distinct advantage because in this way the test has much higher sensitivity and specificity. The PCR method can also be used during the window period. However, PCR is very expensive relative to ELISA.

The main policy that the blood banks use in the USA and some countries in Europe is the following: All blood samples are tested in the ELISA station in batches (due to the lower cost). The batches found clean from this test are re-tested individually in the PCR station. If a blood sample is found clean from both of the stations, then it is stored and ready to be used. Batches found contaminated at ELISA, and items found contaminated at PCR, are discarded.

Every clean sample is separated into three components: red blood cells, platelets and plasma. Each of the components may be used for different purposes and has a different shelf life. Red blood cells have 35 to 42 days shelf life. Plasma has 12 to 15 months shelf life. Platelets only have 4-6 days shelf life, which means that they must be used within 4 to 6 days.
Obviously, the test procedures must be completed as fast as possible due to the limited life time of the components of the blood. This gives rise to a quite complicated optimization problem. On the one hand, high accuracy of the tests is required – hence both the ELISA test and the PCR test are performed. On the other hand, speed is required. The PCR test is time consuming, but one could speed up by having several PCR machines available – which comes at a price.

We develop two blood testing models in this paper. Model I is a model of two queueing systems in series. The first queueing system models the ELISA testing operation as a single server queue in which customers (items) are served (tested) in batches of a fixed size. The second one models the PCR testing operation as a multi-server queue in which customers are served individually. When a batch service is completed in the first queueing system, the batch leaves the whole system with some fixed probability (corresponding to the batch being contaminated), and with the complementarity probability it enters the second queueing system. We present a detailed performance analysis of Model I. Furthermore, we define a profit objective function: mean profits minus mean costs. Profits are made for each blood item that is found clean in the tests, provided the tests took less time than a certain "window size" l. There is also a reward for each unit of time the tests are concluded before l. The costs are screening costs and acquisition costs. The key decision variable is the batch size m; however, in some cases we also take l as an additional decision variable. We use the results of the performance analysis to optimize the profit objective function, determining the optimal choices of batch size m and window size l.
Model II explores an option that is not yet common in CBS's. It is based on the following observation. The ELISA test is done in batches, and is not that expensive. An interesting option is to do a *first* ELISA test with a large batch size m; and if the batch is *not* found clean, one could divide the batch in several smaller batches and test again. Probably most of the smaller batches are found clean, and can still be sent to the PCR. Accordingly, Model II differs from Model I in the following respect: The first of the two queueing systems in series is a single server queue in which batches of size m, after service, are not discarded if they are seen to be contaminated. Instead, such a batch is split into $k = \frac{m}{m_1}$ batches of size $m_1 < m$, and these batches join the end of the queue to receive another service. In principle, when such batches have been tested, they may be split once again, now in batches of size m_2; etc. Just as in Model I, batches that are found clean enter a multi-server queue that represents the PCR testing operation.
While Model I can be and will be analyzed in much detail, Model II is much more involved. Also in view of the fact that the feedback option of Model II is not yet implemented in blood banks, we shall only sketch a possible approach to Model II. We shall propose an approximation approach to Model II, which could be explored in more detail in the future, to assess the possible benefits of the feedback option.

The paper is organized as follows. In Section 2 we describe Model I. Section 3 presents a queueing-theoretic analysis of the model. Section 4 contains numerical results and an optimization study for the model. Section 5 describes Model II, and sketches a way to analyse and optimise that model.

2. Model I: Description

Model I is a model of two queues in series: Q_1 and Q_2. The model is displayed in Figure 1. Q_1 is a single server queue; it represents the ELISA blood testing operation. We

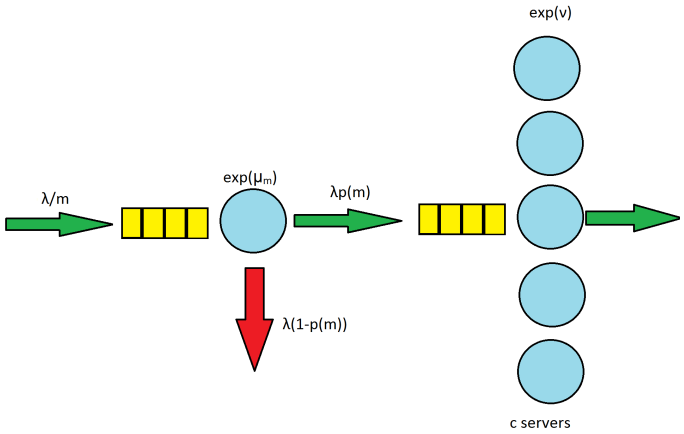

Figure 1. Model I

assume that customers (items) arrive in batches of fixed size m. The arrival process of batches is a Poisson process of rate $\frac{\lambda}{m}$. These batches are served as a whole, one batch after another, in order of arrival. The service time of the ith batch is denoted by B_{1i}, $i = 1, 2, \dots$. Successive service times are independent, and are exponentially distributed with rate $\mu_m = \frac{1}{a_0+a_1 m}$, with a_0, a_1 positive parameters. Hence the mean service time equals $a_0 + a_1 m$, and therefore is linear in the batch size m. It should be noticed that m is going to be a decision variable in our model. The above implies that Q_1 is a simple $M/M/1$ queue.

Assume that an arbitrary item is contaminated with probability ϵ (this is a value that is well known by biostatisticians. In the sequel, we shall typically take $\epsilon = 0.001$; data recently provided by the Israeli Blood Bank for 2011 and 2012 suggest that $\epsilon = 7 \times 10^{-4}$). The probability that a batch of size m is clean hence equals $p(m) := (1-\epsilon)^m$. So a batch of size m is discarded with probability $1 - p(m)$, i.e., leaves the system; and with probability $p(m)$ it enters Q_2 upon departure from Q_1.

It is well known that the departure process of an $M/M/1$ queue is a Poisson process (Burke's output theorem; see, e.g., p. 288-289 of [8]). Furthermore, the splitting of a Poisson process with fixed probabilities also results in (independent) Poisson processes; so, in particular, the arrival process of successive batches at Q_2 is Poisson, with rate $\lambda p(m)/m$. Q_2 represents the PCR blood testing operation. We assume there are c test machines available, and customers are served *individually*, in order of batch arrival; within a batch, the order of service is completely random. The service time of the ith served customer in Q_2 is denoted by B_{2i}, $i = 1, 2, \ldots$. Successive service times are independent, and are exponentially distributed with rate μ. All service times at Q_1 and Q_2 are assumed to be independent, and also independent of the external arrival process. The above implies that Q_2 is an $M^{[m]}/M/c$ queue. Furthermore, Burke [3] has proven that the sojourn times which a customer experiences in an $M/M/1$ queue in series with an $M/M/c$ queue are independent. One of his two proofs uses reversibility of the queue length process at Q_1, and that reversibility argument still holds in our case. This will allow us to handle the total sojourn time of a customer in the model of Q_1 and Q_2 in series, as a sum of two independent sojourn times. In the next section we turn to this analysis.

Remark. Blood units indeed often arrive in batches, but in reality the batch size will be random. This would give rise to a single server queue with batch arrivals of random size and batch services of fixed size. Such a model is tractable (cf. [1]), but (i) the resulting formulas are very complicated, and (ii) one loses the properties that the departure process is Poisson and that sojourn times of a customer in both queues are independent. In this study, we aim to develop an approach which allows one to get useful qualitative and quantitative insight into the effect of particular blood testing procedures, and into the effect of choosing particular batch sizes; for those purposes, it is preferable to start with a relatively simple model. In a future study, we might relax both the assumption of fixed sizes of arriving batches, as well as the assumption of exponentially distributed service times.

3. Model I: Analysis

In this section we mainly focus on the steady-state sojourn time S_1 which an arbitrary batch – or, equivalently, an arbitrary individual customer – experiences in Q_1, and the steady-state sojourn time S_2 which an arbitrary individual customer experiences in Q_2. The distribution of the steady-state sojourn time of a customer, S_1, in the $M/M/1$ queue Q_1 is known to be exponential [9]:

$$P(S_1 > x) = \mathrm{e}^{-(\mu_m - \frac{\lambda}{m})x}, \ x \geq 0. \tag{1}$$

Now let us turn to S_2. Q_2 is an $M^{[m]}/M/c$ queue. Batches of size m arrive at the c-server queue Q_2 according to a Poisson process with rate $\hat{\lambda} = \frac{\lambda}{m} p(m)$. Customers are served one by one, in order of arrival; customers from the same batch are served in random order. The service times at Q_2 are exponentially distributed, with rate μ. An exact analysis of the queue length distribution and delay distribution in the $M^{[X]}/M/c$ queue with batch arrivals of random size X has been provided by Cromie et al. [5]. For our purposes, the approximations that they also provide for these performance measures are more suitable. We follow the discussion of Tijms [16], pp. 304-305; see also Eikeboom and Tijms [6],

who approximate the queue length and sojourn time distributions in the more general $M^{[X]}/G/c$ queue.

Denote by p_j, $j = 0, 1, \ldots$, the steady-state probability of j customers being present in the $M^{[m]}/M/c$ queue Q_2. They satisfy the following balance equations:

$$\begin{aligned} \min(j,c)\mu p_j &= \hat{\lambda} \sum_{i=0}^{j-1} p_i, \ j \le m, \\ \min(j,c)\mu p_j &= \hat{\lambda} \sum_{i=j-m}^{j-1} p_i, \ j > m. \end{aligned} \tag{2}$$

In addition, the p_j satisfy the normalizing condition $\sum_{j=0}^{\infty} p_j = 1$.

Let W denote the steady-state waiting time of an arbitrary customer, and let $W^{(1)}$ denote the steady-state waiting time of the *first* customer of a batch. Then, cf. [5, 6],

$$P(W > x) = \frac{1}{\rho} \sum_{i=0}^{\infty} e^{-c\mu x} \frac{(c\mu x)^i}{i!} \sum_{j=1}^{\infty} p_{c+i+j}, \ x \ge 0, \tag{3}$$

$$P(W^{(1)} > x) = \sum_{i=0}^{\infty} e^{-c\mu x} \frac{(c\mu x)^i}{i!} \sum_{j=0}^{\infty} p_{c+i+j}, \ x \ge 0. \tag{4}$$

Tijms [16], pp. 304-305, see also Cromie et al. [5], suggests an accurate method of approximating p_j and $P(W > x)$ for x large enough. To reduce the computational effort of solving the set of equations (2), he suggests to use the asymptotic expansion

$$p_j \approx \sigma \tau^{-j} \text{ for j large enough,} \tag{5}$$

with τ the unique solution of the equation

$$\hat{\lambda}\tau(1 - \tau^m) = c\mu(1 - \tau), \tag{6}$$

on the interval $(1, \infty)$ and σ being given by

$$\sigma = \frac{(\tau - 1)\sum_{i=0}^{c-1}(c - i)p_i\tau^i/c}{1 - \hat{\lambda}(m - 1)\tau^{m+1}/(c\mu)}. \tag{7}$$

Tijms suggests to use (2) for j smaller than some number K, starting with $p_0 := 1$ and successively computing $p_1, p_2, \ldots, p_{K-1}$, and then using (5) for $j \ge K$. Finally, all p_j are obtained by normalization.

For $P(W > x)$, Tijms [16] suggests the following approximation:

$$P(W > x) \approx \frac{\sigma\tau^{-c}}{\tau - 1} \mathrm{e}^{-c\mu(1-1/\tau)x}, \text{ for x large enough.} \tag{8}$$

Optimization problem

We now turn to our optimization problem. We aim to maximize a particular profit objective function, under certain assumptions on the batch size m and the window size l, i.e., the maximum time we allow a blood item to spend in the system so that it can still be used effectively.

In order to define the objective function, we introduce various cost and profit parameters:

(1) $k_e(m) > 0$ represents the cost for screening a group of size m at the ELISA station. This cost is an increasing function of m. We assume that $k_e(m) = 1 + \frac{m-1}{16}$. Measurements at the Israeli CBB tend to give this relation between the cost in the ELISA station and the size of the group. The size of the batch has an upper bound m_0 due to the limited equipment of the station. A realistic value of m_0 is 48.
(2) $k_{pcr} > 0$ is the cost for screening a single item at the PCR station. The screening cost at the PCR station is much higher than the one at the ELISA station. Therefore, k_{pcr} satisfies the relation $k_{pcr} > k_e(m_0)$.
(3) $l > 0$ represents an upper bound of the residual shelf life. The shelf life of the blood units is quite limited (the components of the blood, such as platelets, can not endure for a long time). l can also be considered a decision parameter provided with an upper bound, say $l_0 = 96$ hours.
(4) $r > 0$ is a reward for each time unit of residual shelf life of a single item.
(5) $r_1 > 0$ is a reward for each item found clean in both of the stations, provided the total time it has spent in the system is smaller than l.
(6) $d > 0$ represents the acquisition cost per item.

Finally we remind the reader that it is possible that a unit is found contaminated in the PCR station, although it has been declared clean in the ELISA station. The corresponding conditional probability is denoted by $\gamma \in (0, 1)$. A typical value is $\gamma = 5 \times 10^{-5}$ (based on data provided by the Israeli CBS). We shall consider the following profit objective function:

$$\begin{aligned} R(m) &= r\lambda p(m)(1-\gamma)E[(l - S_1 - S_2)I(l - S_1 - S_2 > 0)] \\ &+ r_1\lambda p(m)(1-\gamma)P(l - S_1 - S_2 > 0) - k_{pcr}\lambda p(m) - k_e(m)\lambda/m - d\lambda, \end{aligned} \tag{9}$$

under the assumptions: $m \le m_0$, $l \le l_0$. Here I denotes an indicator function. In some cases we shall also explicitly take the dependence of R on l into account, and write $R(m, l)$. We briefly explain the rationale behind this formula. The first term represents the expected reward for each time unit of residual shelf life of an item provided the unit is found clean in both stations. The second term represents the expected reward for each clean item provided its residual shelf life is less than l. The third term represents the cost for the PCR test. The fourth term represents the cost for the ELISA test and finally the fifth term is the acquisition cost of an item.

As we can see, the profit objective function contains the probability $P(l - S_1 - S_2 > 0) = P(S_1 + S_2 < l)$ and the expectation $E[(l - S_1 - S_2)I(l - S_1 - S_2 > 0)]$. We will try to find analytical expressions for these two terms, which both involve the total sojourn time $S_1 + S_2$. We can write $S_2 = W_2 + B_2$, where W_2 denotes an arbitrary waiting time and B_2 an arbitrary service time in Q_2 (which is exponentially distributed with rate μ); W_2 and B_2 are independent. Observing that $S_1 \sim \exp(a)$ with $a = \mu_m - \frac{\lambda}{m}$, cf. (1), and that $B_2 \sim \exp(\mu)$, and assuming for the moment that $P(W_2 > x) = \zeta e^{-\delta x}$, with $\zeta = \frac{\sigma\tau - c}{\tau - 1}$ and $\delta = c\mu(1 - 1/\tau)$, i.e., assuming that (8) holds for all $x \ge 0$, one would obtain the following result (where, for expository reasons, we have taken $\lambda_1 = \mu_m - \frac{\lambda}{m}$, $\lambda_2 = \mu$ and $\lambda_3 = \delta$):

$$P(S_1+S_2 > l) = (1-\zeta)\left(\frac{\lambda_1}{\lambda_1 - \lambda_2}e^{-\lambda_2 l} + \frac{\lambda_2}{\lambda_2 - \lambda_1}e^{-\lambda_1 l}\right) + \zeta\sum_{i=1}^{3}\prod_{j\neq i}\frac{\lambda_j}{\lambda_j - \lambda_i}e^{-\lambda_i l},\ l \ge 0. \tag{10}$$

For large values of l, (10) is a good approximation, because the distributions of S_1 and B_2 are exact while the approximation in (8) is good for large x. For small values of l, (10) is not a good approximation, because the approximation in (8) is in general not good for small values of l. However, we are mainly interested in quite large values of l, and then $P(S_1+S_2>l)$ typically is very close to 0; and even a 50% error in that probability translates into a very small error in $P(l-S_1-S_2>0)$.

We now proceed to the calculation of the expectation $E[(l-S_1-S_2)I(l-S_1-S_2>0)] = E[(l-S_1-B_2-W_2)I(l-S_1-B_2-W_2>0)]$. Let us define $g(S_1,B_2,W_2) = (l-S_1-B_2-W_2)I(l-S_1-B_2-W_2>0)$. So, we are looking for the $E[g(S_1,B_2,W_2)]$. Distinguishing between $W_2=0$ and $W_2>0$, we have

$$E[g(S_1,B_2,W_2)] = (1-\zeta)E[g(S_1,B_2,0)] + \zeta E[g(S_1,B_2,W_2)|W_2>0]. \tag{11}$$

We now calculate these two summands.

• First summand: The variables S_1 and B_2 are exponential with parameters λ_1 and λ_2 respectively. Therefore,

$$\begin{aligned} E[g(S_1,B_2,0)] &= \int_0^\infty\int_0^\infty (l-x-y)I(l-x-y>0)\lambda_1 e^{-\lambda_1 x}\lambda_2 e^{-\lambda_2 y}dxdy \\ &= \int_0^l \lambda_2 e^{-\lambda_2 y}\left[\int_0^{l-y}(l-x-y)\lambda_1 e^{-\lambda_1 x}dx\right]dy \\ &= \left(\frac{1}{\lambda_1}-\frac{1}{\lambda_1-\lambda_2}\right)e^{-\lambda_1 l} + \left(\frac{1}{\lambda_2}-\frac{1}{\lambda_2-\lambda_1}\right)e^{-\lambda_2 l} + \left(l-\frac{1}{\lambda_1}-\frac{1}{\lambda_2}\right). \end{aligned} \tag{12}$$

• Second summand: $W_2|W_2>0$ is exponential with parameter λ_3. Using this fact, we can write the expectation as

$$\begin{aligned} & E[g(S_1,B_2,W_2)|W_2>0] \\ = & \int_0^\infty\int_0^\infty\int_0^\infty (l-x-y-z)I(l-x-y-z>0)\lambda_1 e^{-\lambda_1 x}\lambda_2 e^{-\lambda_2 y}\lambda_3 e^{-\lambda_3 z}dxdydz \\ = & \int_0^l \lambda_3 e^{-\lambda_3 z}\int_0^{l-z}\lambda_2 e^{-\lambda_2 y}\int_0^{l-y-z}\lambda_1 e^{-\lambda_1 x}(l-x-y-z)dxdydz \\ = & \frac{\lambda_2\lambda_3}{\lambda_1(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)}e^{-\lambda_1 l} + \frac{\lambda_1\lambda_3}{\lambda_2(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}e^{-\lambda_2 l} \\ + & \frac{\lambda_1\lambda_2}{\lambda_3(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}e^{-\lambda_3 l} + l - \frac{1}{\lambda_1}-\frac{1}{\lambda_2}-\frac{1}{\lambda_3}. \end{aligned} \tag{13}$$

Substitution of (12) and (13) in (11) yields $E[g(S_1,B_2,W_2)] = E[(l-S_1-S_2)I(l-S_1-S_2>0)]$.

4. Model I: numerical results

In this section we consider the profit objective function $R(m)$ given in (9) for a particular choice of the model parameters, and for various values of our decision variables: the batch size m and the window size for the shelf life, l. For the other parameters we have taken values which are based on data which were kindly supplied to us by Professor Shinar, director of the Israeli bloodbank. In each of the first five tables and the plots, these values are the same (unless indicated otherwise). These values are: $\lambda = 2$, $\mu_m = \frac{1}{0.079m+1.921}$ hours^{-1}, $\nu = \frac{1}{6}$ hours^{-1}, $\epsilon = 10^{-3}$, $\gamma = 5\times 10^{-5}$, $c = 20$, $k_e(m) = 1+\frac{m-1}{16}$, $k_{pcr} = 5$,

$d = 1$, $r = 0.1$, $r_1 = 3$.

In Table 1 we print the profit objective function R for $m = 5(1)48$, keeping l fixed at 72 hours. The optimal choice for m appears to be $m = 12$ and the corresponding mean profit value is $R(12) = 5.5073$. We refer to Section 2.5 of [17] for a discussion of the accuracy of the approximations we use for $P(S_1 + S_2 > l)$ and $E[(l - S_1 - S_2)I(l - S_1 - S_2 > 0)]$, which are both based on approximation (8). The conclusion in [17] was that the error in the former term is almost negligible, and that the error in the latter term is in most tested cases below 1%.

TABLE 1. Results for the profit objective function R, for $l = 72$.

m	R	m	R	m	R	m	R
5	0.4693	16	5.4429	27	4.9714	38	4.3991
6	4.2636	17	5.4114	28	4.9210	39	4.3464
7	4.9731	18	5.3763	29	4.8700	40	4.2937
8	5.2522	19	5.3383	30	4.8186	41	4.2410
9	5.3892	20	5.2979	31	4.7669	42	4.1886
10	5.4603	21	5.2554	32	4.7148	43	4.1362
11	5.4951	22	5.2111	33	4.6625	44	4.0841
12	5.5073	23	5.1653	34	4.6100	45	4.0322
13	5.5045	24	5.1183	35	4.5574	46	3.9805
14	5.4910	25	5.0701	36	4.5047	47	3.9291
15	5.4699	26	5.0212	37	4.4519	48	3.8779

In Figure 2 we also display the graph of $R(m)$ for this same case of $l = 72$. In Table 2 we

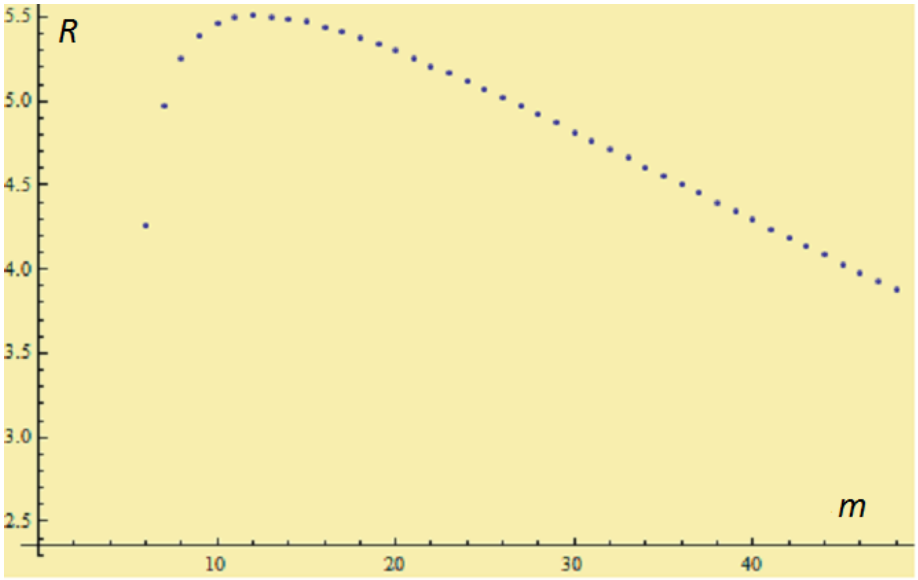

FIGURE 2. Profit objective function $R(m)$ for the case $l = 72$

let λ and c vary, again taking $l = 72$. We indicate which value of $m \in \{1, 2, \ldots, 48\}$ yields the largest profit, and we give that profit.

TABLE 2. Optimal choices for m.

λ	c	optimal m	$R(m)$
1	10	7	2.8379
1.5	10	5	2.0635
1	15	10	3.0000
2	15	9	4.5273
1	20	11	3.0253
2	20	12	5.5073
3	20	13	3.6650
1	25	11	3.0304
2	25	14	5.6874
3	25	16	7.0838
3.5	25	17	5.9829
1	30	11	3.0314
2	30	15	5.7405
3	30	19	7.6863
4	30	22	7.1480
4.5	30	24	3.4600

In Table 3 and Figure 3 we print R for a wide range of (m, l) combinations. In the table, $-$ denotes that the stability condition ($\rho_1 < 1$ and $\rho_2 < 1$) is violated. In Table 4, λ and c are being varied. For each of 27 (λ, c) combinations, we display the R value that is largest among all (m, l) combinations with $m = 4(2)48$ and $l = 12(6)96$. It turns out to be optimal to take l equal to the maximal value of 96 hours.
In Table 5 we consider exactly the same cases as in Table 4, except that we restrict l to values 12(6)48.
From all these instances, we see that this monotonicity of the function R is repeated in different cases and for different values of the parameters. The conclusion of all these trials is that the function R is generally increasing for smaller values of m in the interval [1,48] until it reaches a local maximum and then it decreases as m becomes larger. That is why we can actually find a maximal value of R in a particular point m provided l is fixed. Generally speaking, increasing the value of m provides a benefit for the ELISA station and a disadvantage for the PCR station. If the size of the group increases, then the testing process in the ELISA stations is executed faster for more donors. But once this group arrives at the PCR station, then there are more donors arriving simultaneously and that fact increases the waiting time in the PCR station.

Remark. If the service times at the PCR queue are not exponentially distributed, then one may proceed as follows. First consider the case that service times are deterministic (denoted by D). For that case, Eikeboom and Tijms [6] (see also Tijms [16], pp. 305-308), present an approximation for $P(W > x)$ that is similar in nature to (8). That

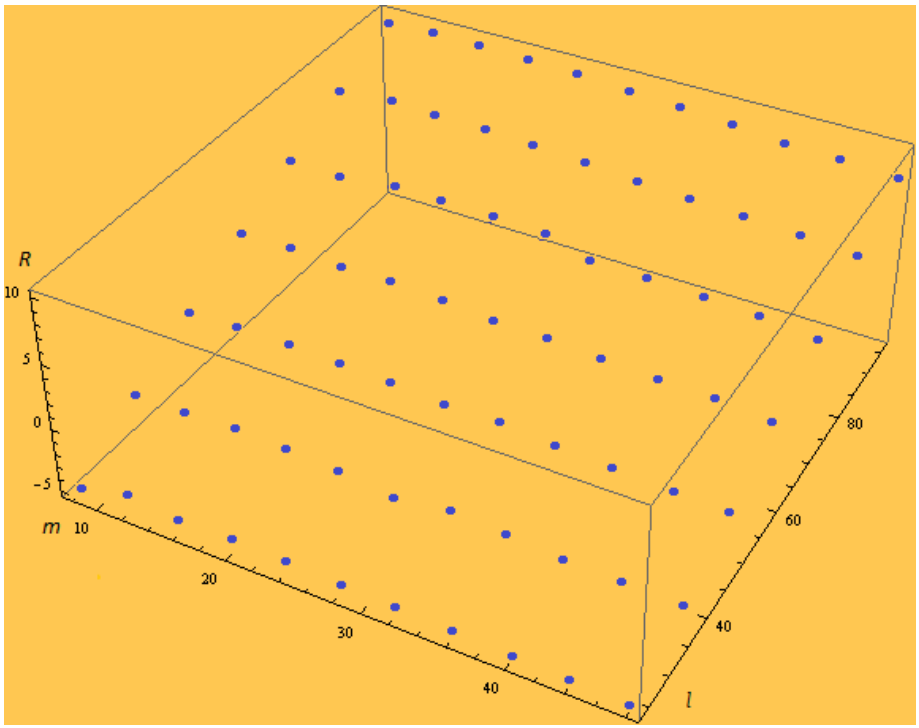

FIGURE 3. Profit objective function $R(m,l)$ for a range of (m,l) combinations

approximation has been implemented in [17], leading to the numerical results contained in Table 6 (see Tables 53 and 55 of [17] for further numerical experiments). This time we take the following parameter values: $\lambda = 0.5$, $\mu_m = \frac{1}{0.079m+1.921}$ hours^{-1}, $D = 6$ hours, $\epsilon = 10^{-3}$, $\gamma = 5 \times 10^{-5}$, $c = 30$, $l = 72$ hours, $k_e(m) = 1 + \frac{m-1}{16}$, $k_{pcr} = 5$, $k_{ser} = 0.5$, $d = 1$, $r = 0.3$, $r_1 = 5$.

We denote by E the expectation $E[(l-S_1-S_2)I(l-S_1-S_2>0)]$ and by P the probability $P(l-S_1-S_2>0)$. In addition the terms of cost are denoted by $cost_1 = -k_{pcr}\lambda p(m) - k_e(m)\lambda/m - d\lambda$. Among the values of m that we have chosen above, it seems that the optimal one is $m = 8$. Let us examine the results. The occupation rates are getting smaller while m increases because a larger m implies a smaller arrival rate at ELISA and therefore at PCR too. The sojourn time $E(S_1)$ increases as it was expected while $E(S_2)$ follows a different behavior. There is a similarity between the results we get here and the ones that we have obtained above for the case in which the service times in the PCR station are exponential.

Finally, if the service times are generally distributed, one may resort to the approximation for the waiting time distribution which is proposed in [6], see also [16], p. 309. That approximation is based on the approximations for $M/M/c$ and $M/D/c$, interpolating the corresponding waiting time percentiles for those two models.

5. Model II: Model description, and a sketch of its analysis

Just like Model I, Model II is a model of two queues in series: Q_1 and Q_2. We first describe Q_1 in detail. Q_1 is a single server queue; it represents the ELISA blood testing

TABLE 3. Results for the profit objective function.

m	l	R	m	l	R	m	l	R
4	24	-	20	48	0.5936	36	72	4.5047
8	24	-4.6406	24	48	0.4384	40	72	4.2937
12	24	-4.1920	28	48	0.2686	44	72	4.0841
16	24	-4.1618	32	48	0.0932	48	72	3.8779
20	24	-4.2280	36	48	-0.0827	4	84	-
24	24	-4.3317	40	48	-0.2566	8	84	7.6335
28	24	-4.4538	44	48	-0.4267	12	84	7.8786
32	24	-4.5856	48	48	-0.5917	16	84	7.8047
36	24	-4.7224	4	60	-	20	84	7.6501
40	24	-4.8615	8	60	2.8687	24	84	7.4608
44	24	-5.0006	12	60	3.1353	28	84	7.2534
48	24	-5.1380	16	60	3.0808	32	84	7.0366
4	36	-	20	60	2.9458	36	84	6.8151
8	36	-1.9721	24	60	2.7769	40	84	6.5917
12	36	-1.6423	28	60	2.5912	44	84	6.3688
16	36	-1.6642	32	60	2.3979	48	84	6.1484
20	36	-1.7672	36	60	2.2022	4	96	-
24	36	-1.9006	40	60	2.0070	8	96	10.0144
28	36	-2.6466	44	60	1.8145	12	96	10.2499
32	36	-2.1969	48	60	1.6266	16	96	10.1665
36	36	-2.3474	4	72	-	20	96	10.0024
40	36	-2.4965	8	72	5.2522	24	96	9.8036
44	36	-2.6428	12	72	5.3073	28	96	9.5867
48	36	-2.7855	16	72	5.4429	32	96	9.3600
4	48	-	20	72	5.2979	36	96	9.1284
8	48	0.4741	24	72	5.1183	40	96	8.8944
12	48	0.7592	28	72	4.9210	44	96	8.6604
16	48	0.7165	32	72	4.7148	48	96	8.4281

operation. We assume that customers (items) arrive in batches of fixed size $m = m_0$. The arrival process of batches is a Poisson process of rate $\lambda_0 = \frac{\hat{\lambda}}{m_0}$, where $\hat{\lambda}$ is the arrival rate of blood items at the ELISA station. A batch of size m_0 is clean with probability $(1-\epsilon)^{m_0} \approx 1-m_0\epsilon$, which for realistic values like $m_0 = 48$ and $\epsilon = 7\times10^{-4}$ is approximately 0.97. Such a clean batch enters Q_2 (the PCR station). However, if a batch of size m_0 is found contaminated then, unlike in Model I, it is not scrapped but split in k_1 batches of size $m_1 = m_0/k_1$. These batches are instantaneously resent to the queue of Q_1. When all these k_1 batches of size m_1 have been tested (served), at least one of them will definitely *not* be clean. When a batch of size m_1 is found clean, it will enter Q_2. When a batch of size m_1 is found contaminated, then it may also not be scrapped but split in k_2 batches of size $m_2 = m_1/k_2$. These batches are instantaneously resent to the queue of Q_1. This recycling process may go on for several cycles, the ith recycle having batches of size m_i,

TABLE 4. Optimal choices for m and l.

c	λ	optimal m	optimal l	R
10	1	6	96	5.2191
10	1.5	6	96	5.5576
15	1	10	96	5.3760
15	1.5	10	96	7.7839
15	2	10	96	9.2516
15	2.5	48	96	-5.9564
20	1	10	96	5.4003
20	1.5	12	48	7.9490
20	2	12	96	10.2499
20	2.5	12	96	11.8769
20	3	12	96	10.6075
25	1	10	96	5.4044
25	1.5	12	96	7.9874
25	2	14	96	10.4205
25	2.5	14	96	12.5760
25	3	16	96	14.1689
25	3.5	18	96	14.1787
25	4	18	96	6.3998
30	1	10	96	5.4051
30	1.5	12	96	7.9963
30	2	14	96	10.4706
30	2.5	16	96	12.7635
30	3	18	96	14.7564
30	3.5	20	96	16.2067
30	4	22	96	16.5122
30	4.5	24	96	13.5700
30	5	48	96	-4.7617

$i = 1, 2, \ldots$. In principle, one will most likely not do more than two recycles; but let us assume, generally, that there are at most r recycles.

Before discussing the distribution of an arbitrary batch, let us first specify the service time distribution of a batch. We assume that all service times are independent (and also independent of interarrival times), and exponentially distributed with mean $\mu_m = a_0 + a_1 m$ when the batch size is m, with a_0, a_1 positive parameters. The order of service is FCFS: service in order of arrival, and when a batch is split into k_i batches, these k_i batches join the end of the queue in an arbitrary order. So, just like in Model I, the mean service time in Q_1 is linear in the batch size m: the mean service time of a batch of items which have already been resent j times equals $a_0 + a_1 m_j$. m_0 is again going to be a decision variable in our model, but now $m_1, m_2, \ldots$ are also decision variables.

One can easily determine the probability, $s_i := P(X = m_i)$, that an arbitrary batch is of size $X = m_i$; i.e., the fraction of batches that is in (re-)cycle i. Let Y_i denote the number

TABLE 5. Optimal choices for m and l.

c	λ	optimal m	optimal l	R
10	1	6	48	0.4469
10	1.5	6	48	-1.2987
15	1	10	48	-0.6223
15	1.5	10	48	0.6538
15	2	10	48	-0.2080
15	2.5	48	48	-16.9131
20	1	12	48	0.6501
20	1.5	12	48	0.8316
20	2	12	48	0.7592
20	2.5	14	48	0.0358
20	3	14	48	-2.9662
25	1	12	48	0.6564
25	1.5	14	48	0.8742
25	2	14	48	0.9469
25	2.5	16	48	0.7544
25	3	18	48	-0.0004
25	3.5	18	48	-2.0438
25	4	18	48	-9.0935
30	1	12	48	0.6577
30	1.5	14	48	0.8870
30	2	16	48	1.0062
30	2.5	18	48	0.9496
30	3	20	48	0.5993
30	3.5	22	48	-0.2648
30	4	24	48	-2.0898
30	4.5	24	48	-6.0973
30	5	48	48	-18.4884

TABLE 6. Results for the profit function when PCR service times are constant.

m	E	P	$r\lambda p(m)(1-\gamma)E$	$r_1\lambda p(m)(1-\gamma)P$	$cost_1$	R
4	61.7806	1	9.2296	2.4900	-2.1425	9.5772
8	62.1032	1	9.2408	2.4801	-2.0779	9.6429
16	61.9409	1	9.1431	2.4603	-2.0367	9.5667
24	61.4125	1	8.9929	2.4407	-2.0152	9.4184
48	57.2840	0.9999	8.1893	3.3824	-1.9707	8.6664

of contaminated items in an arbitrary batch of size m_i, $i = 0, 1, \ldots$. Then it is readily

seen that, for $i = 0, 1, \ldots, r$,

$$P(X = m_i) = \frac{\prod_{j=1}^{i} k_j P(Y_i = 0, Y_{i-1} > 0, \ldots, Y_0 > 0)}{P(Y_0 = 0) + \sum_{h=1}^{r} \prod_{j=1}^{h} k_j P(Y_h = 0, Y_{h-1} > 0, \ldots, Y_0 > 0)}. \tag{14}$$

Sketch of the performance analysis of Model II

Contrary to Model I, which is a simple $M/M/1$ queue, we are now faced with a quite complicated feedback queue. In [2] the authors study an $M/M/1$ queue with feedback probability $p(i)$ of customers who have already received i services. For that model the steady-state joint distribution is obtained of the numbers of customers in their first, second, ... visit. However, that model does not allow batches and, more importantly, the exponential service rate of every customer class (i.e., in its first, second, ... visit) is *the same*. The latter property gives rise to a product form for the joint queue length distribution of all classes. For the present model, with batch services and unequal service rates, it will be very challenging to obtain the steady-state joint queue length distribution, and subsequently the sojourn time distribution of customers (batches), and one seems forced to take recourse to approximations.

We suggest to handle Model II by treating Q_1 as an $M/G/1$ queue with $r + 1$ customer classes. Class 0 consists of the newly arriving customers (batches) which are clean and do not require any recycle. Their service times are exponentially distributed, with mean $a_0 + a_1 m$. Class 1 consists of the batches of size m_1 which have first received an exponential service of mean $a_0 + a_1 m$, and subsequently an exponentially distributed service time with mean $a_0 + a_1 m_1$, and who thereafter leave Q_1. Classes $2, \ldots, r$ are similarly defined, class i having service times which are the sum of $i + 1$ independent, exponentially distributed times. Assuming (our *first* approximation assumption) that all customer classes arrive according to independent Poisson processes with appropriate rates ($\lambda\times$ the probability that an arbitrary customer is of type i), and that all customers are served FCFS, all customer types have the same waiting time distribution, which is given by the familiar Pollaczek-Khintchine distribution. The sojourn time distribution immediately follows via a convolution of the waiting time distribution with the service time distribution.

In reality the feedback flows are *not* Poisson. However, the feedback probability of a type-0 batch, $1 - p(m) = 1 - (1 - \epsilon)^m$, is very small. Typical parameter values are $\epsilon = 7 \times 10^{-4}$ and $m = 48$, yielding $1 - p(m) \approx m\epsilon = 3,36 \times 10^{-2}$. Hence class-0 customers are dominant in Q_1, and feedback is quite rare: approximately 1 out of 30 in the above case $m = 48$, and approximately 1 out of 60 if m were 24. It should be noted that type-1 customers have a relatively high feedback probability, since at least one of the m customers in the batch is contaminated, implying that at least one of the $k_1 = m/m_1$ batches, in which the original batch is split, will be contaminated. Still, type-2 customers are even more rare than type-1 customers, and so on.

A second complication, w.r.t. the analysis of Model I, is that the output process of Q_1 no longer is a Poisson process – and hence neither is the input process of Q_2. Again, the low feedback probability of newly arriving batches implies that Q_1 behaves quite similarly to an $M/M/1$ queue, and hence the output process of Q_1 is quite accurately approximated by a Poisson process (our *second* approximation assumption).

Let us now describe Q_2 in detail. Q_2 represents the PCR blood testing operation. We assume there are c test machines available, and customers are served *individually*, in order of

batch arrival; within a batch, the order of service is completely random. Successive service times are independent, and are exponentially distributed with rate μ. All service times at Q_1 and Q_2 are assumed to be independent, and also independent of the external arrival process. The above implies that Q_2 is a $G^{[X]}/M/c$ queue; and the second approximation assumption above says that we can view it as an $M^{[X]}/M/c$ queue.
Using the same profit objective function as for Model I, cf. (9), we need to determine the distribution of the sum of the sojourn times S_1 and S_2 of a customer (blood unit) in Q_1 and Q_2. In Model I, those two sojourn times are independent, because Q_1 is an $M/M/1$ queue. Once more, the fact that Q_1 in Model II behaves very much like an $M/M/1$ queue allows us to assume that S_2 is (almost) independent of S_1 (our *third* approximation assumption). Finally, we need to determine the distribution of sojourn time S_2 in Q_2. For this we can use the results of Eikeboom and Tijms [6]; in particular, we can again use the sojourn time expressions for S_2 which we used for Model I in Section 3 (our *fourth* approximation assumption). If the service times in Q_2 are generally distributed, we could use the sojourn time approximation which was presented in [6]. Some preliminary results, which suggest that this approximation method has good potential, are contained in [17].

Acknowledgment
The authors gratefully acknowledge discussions with Professor Eilat Shinar, director of the Israeli Central Blood Bank. The research of Shaul Bar-Lev and David Perry was supported in part by grant No. I-1184-31.4/2012 from the German-Israel Science Foundation, and grant No. 1071/14 from the Israel Science Foundation. The research of Onno Boxma was supported by the NETWORKS project, funded by the Dutch government.

References

[1] T.P. Bagchi and J.G.C. Templeton (1972). Numerical Methods in Markov Chains and Bulk Queues. Springer-Verlag, Berlin.
[2] J.L. van den Berg and O.J. Boxma (1991). The $M/G/1$ queue with processor sharing and its relation to a feedback queue. *Queueing Systems* 9: 365-402.
[3] P.J. Burke (1968). The output process of a stationary $M/M/s$ queueing system. *Ann. Math. Statist.* 39: 1144-1152.
[4] J.A. Chiavetta, M. Escobar, A.M. Newman, Y. He, P. Driezen, S. Deeks, D.E. Hone, S.F. O'Brien and G. Sher (2003). Incidence and estimated rates of residual risk for HIV, hepatitis C, hepatitis B and human T-cell lymphotropic viruses in blood donors in Canada, 1990–2000. *CMAJ*, 169: 767-773.
[5] M. V. Cromie, M.L. Chaudry, and W.K. Grassmann (1979). Further results for the queueing system $M^X|M|c$. *J. Oper. Res. Soc.* 40: 755-763.
[6] A. M. Eikeboom and H. C. Tijms (1987). Waiting-time percentiles in the multi-server $M^X|G|c$ queue with batch arrivals. *Probability in the Engineering and Informational Sciences* 1: 75-96.
[7] P. Ghandforoush and T.K. Sen (2010). A DSS to manage platelet production supply chain for regional blood centers. *Decision Support Systems*, 50: 32-42.
[8] M. Harchol-Balter (2013). *Performance Modeling and Design of Computer Systems*. Cambridge University Press, Cambridge.
[9] M. Haviv (2013). *Queues*. Springer-Verlag, Berlin.
[10] M.K. Hourfar, C. Jork, V. Schottstedt, M. Weber-Schehl, V. Brixner, M.P. Busch, G. Geusendam, K. Gubbe, C. Mahnhardt, U. Mayr-Wohlfart, L. Pichl, W.K. Roth, M. Schmidt, E. Seifried and D.J. Wright (2008). Experience of German Red Cross blood donor services with nucleic acid testing: results of screening more than 30 million blood donations for human immunodeficiency virus-1, hepatitis C virus, and hepatitis B virus. *Transfusion*, 48(8): 1558-1566.

[11] B.R. Jackson, M.P. Busch, S.L. Stramer and J.P. AuBuchon (2003). The cost-effectiveness of NAT for HIV, HCV, and HBV in whole-blood donations. *Transfusion*, 43: 721-729.

[12] D.A. Marshall, S.H. Kleinman, J.B. Wong, J.P. AuBuchon, D.T. Grima, N.A. Kulin and M.C. Weinstein (2004). Cost-effectiveness of nucleic acid test screening of volunteer blood donations for hepatitis B, hepatitis C and human immunodeficiency virus in the United States. *Vox Sanguinis*, 86: 28-40.

[13] V. Schottstedt, W. Tuma, G. Bünger and H. Lefèvre (1998). PCR for HCV and HIV-1 experiences and first results from routine screening programme in a large blood transfusion service. *Biologicals*, 26: 101-104.

[14] S.L. Stramer, S.A. Glynn, S.H. Kleinman, D.M. Strong, S. Caglioti, D.J. Wright, R.Y. Dodd and M.P. Busch (2004). Detection of HIV-1 and HCV infections among antibody-negative blood donors by nucleic acid-amplification testing. *The New England Journal of Medicine*, 351(8): 760-768.

[15] S.L. Stramer, U. Wend, D. Candotti, G.A. Foster, F.B. Hollinger, R.Y. Dodd, J.-P. Allain and W. Gerlich (2011). Nucleic acid testing to detect HBV infection in blood donors. *New England Journal of Medicine*, 364: 236-247.

[16] H.C. Tijms (1994). *Stochastic Models – An Algorithmic Approach.* Wiley, New York.

[17] L.P. Vastazos (2014). Queueing Models in Blood Testing Procedures. MSc Thesis, Eindhoven University of Technology.