A novel approach to detect test-seeking behaviour in the blood donor population: making the invisible visible


Transfusion Technology Assessment Unit, Sanquin Research, Amsterdam, the Netherlands
Department of Blood-Borne Infections, Sanquin Research, Amsterdam, the Netherlands
Institute of Mathematics, Astrophysics and Particle Physics, Radboud University, Nijmegen, the Netherlands
Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, the Netherlands

Background and Objectives Individuals may donate blood in order to determine their infection status after exposure to an increased infection risk. Such test-seeking behaviour decreases transfusion safety. Instances of test seeking are difficult to substantiate as donors are unlikely to admit to such behaviour. However, manifestation in a population of repeat donors may be determined using statistical inference.

Materials and Methods Test-seeking donors would be highly motivated to donate following infection risk, influencing the timing of their donation. Donation intervals within 2005–2014 of all Dutch blood donors who acquired syphilis (N = 50), HIV (N = 13), HTLV (N = 4) or HCV (N = 2) were compared to donation intervals of uninfected blood donors (N = 7,327,836) using the Anderson–Darling test. We adjusted for length bias as well as for age, gender and donation type of the infected. Additionally, the power of the proposed method was investigated by simulation.

Results Among the Dutch donors who acquired infection, we found only a non-significant overrepresentation of short donation intervals (P = 0.54). However, we show by simulation that both relatively short and long donation intervals among infected donors can reveal test seeking. The power of the method is >90% if among 69 infected donors >35 (51%) are test seeking, or if among 320 infected donors >90 (30%) are test seeking.

Conclusion We show how statistical analysis may be used to reveal the extent of test seeking in repeat blood donor populations. In the Dutch setting, indications for test-seeking behaviour were not statistically significant. This may, however, be due to the low number of infected individuals.

Key words: donation intervals, repeat donors, test-seeking behaviour.

Introduction

Rather than out of purely altruistic reasons, some individuals may be driven to donate in order to have their infection status determined [1–6]. As this motivation would be strongest following risk behaviour for acquiring infection, test-seeking behaviour poses a threat for blood transfusion safety [7].

It is difficult to identify test-seeking donors as they are likely to be non-compliant with the donor selection questionnaire; they would withhold any concerns about their infection risk in order to avoid deferral. Test-seeking behaviour has been revealed using consequence-free questionnaires in addition to the donor selection process [8]. In various questionnaire studies between 1% and 7% of all donors [1, 8], and between 14% and 54% of all infected donors admitted to being motivated to donate – at least in...
part – to obtain test results [4]. However, donors partaking in such studies might still fear consequences, constraining their truthful response.

Risk behaviour is also frequently identified during intensive post-positive-test counselling interviews [9, 10]. Of counselled Dutch repeat donors, 28% were found to have been non-compliant with the donor health questionnaire, while 42% did not know of, or did not admit to any particular risk factors [10]. As instances of test-seeking behaviour are likely to remain unrevealed even when applying the best questioning techniques, it is difficult to assess the extent of test-seeking behaviour within donor populations.

Despite the lack of reliable information on test-seeking behaviour from individual donors, statistical inference can be used to identify such behaviour for the group of infected repeat donors. If test seeking is present, a comparison of the distribution of interdonation intervals of the infected donors and non-infected donors should reveal differences in donation intervals, hence making the invisible visible.

We will first illustrate our method by testing for test-seeking behaviour in the Dutch repeat donor population. Next, we perform a simulation study to show the ability of the method to detect various levels of test-seeking behaviour in repeat donor populations.

Methods

We define test-seeking donors as those that run an increased risk of infection and who donate in order to acquire test results. As they are highly motivated to know their infection status, they are expected to return for donation more quickly than donors in general. We define these as test-seeking (TS) donors. Yet another group of test-seeking donors might return after a relatively long period of abstention from donating: inactive donors that would not have returned to donate at all, except that they are now motivated to return to obtain test results after engaging in risky behaviour. These latter we define as latent test-seeking (LTS) donors.

Although non-test seekers may also become infected, TS and LTS donors are much more likely to test positive. Therefore, the proportion of test-seeking individuals is expected to be higher among the infected donors as compared to among the uninfected donors. Consequently, TS and LTS behaviour will cause the interdonation intervals of the infected donors to differ from those of non-infected donors.

The distribution of donation intervals for infected donors cannot be compared directly to this distribution for uninfected donors, as the likelihood of becoming infected accumulates with the time between donations. Therefore, the expected distribution of observed donation intervals of infected donors will differ from the distribution of all donation intervals, even in case no test seeking has occurred. More precisely, the distribution of donation intervals of infected donations by non-TS is expected to follow the ‘length-biased’ distribution of all donation intervals [11]. This implies that the occurrence probability of each donation interval is proportional to the length of the donation interval.

To determine whether the donation behaviour of infected donors differs from that of non-infected donors, we first create a reference distribution, which the infected should follow if there were no TS or LTS donors present. A large \( N = 10 \, 000 \, 000 \) reference set represents the length-biased distribution of all donation intervals, as it is created by drawing randomly from all the intervals in accordance with interval length. The Anderson–Darling (AD) test is then applied to test for differences between this reference distribution and that of the observed donation intervals of the infected [12].

Various donor characteristics, for example age and gender of the donor, may be related to both the length of the donation interval and the probability of becoming infected. By ensuring that for our reference set donation intervals are drawn from donors with comparable attributes as the infected donors, we adjust for such potential confounders.

Analysing infections among Dutch donors

Donation intervals of Dutch blood donors who acquired syphilis, HIV, human T-lymphotropic virus (HTLV) or hepatitis C virus (HCV) within 2005–2014 (\( N = 69 \)) were compared to those of all blood donors (\( N = 7 \, 327 \, 836 \) donation intervals of 591 702 donors, representing 19 456 person-years of follow-up). From our main analysis, we excluded:

1. All hepatitis B virus (HBV)-infected donors (\( N = 32 \)), because HBV infection can be cleared naturally. To obtain a corrected reference set for these donors, additional assumptions concerning the duration of testing HBV positive are required.
2. Infected donors with their last prepositive test donation before 2005 (\( N = 29 \)), to ensure a fair comparison with the uninfected donors (for whom information was only available from 2005 and onwards).

In the Supporting information, we describe two more complex analyses in which all infected donors are included. As gender, age (categorized per year) and donation type at last negative donation (whole blood, thrombocytes or plasma) are considered potential confounders, the

© 2016 International Society of Blood Transfusion

Vox Sanguinis (2016) 111, 274–280
combination of each of these factors was added to our reference set \( (N = 10\ 000\ 000) \) in an equal proportion as occurring in the set of infected donors (Fig. 1).

**Simulation methods**

The power of the proposed method to detect test seeking was investigated by simulation. We simulated donation timing of non-test-seeking (non-TS), test-seeking (TS) and latent test-seeking (LTS) donors. TS donors were assigned higher probabilities (per unit time) of becoming infected and higher probabilities to donate. The LTS donors represent otherwise inactive donors who donate only after ‘activation’, which was set to coincide with an increase in their infection risk.

For non-TS donors, we drew intervals randomly from the donation intervals of all Dutch repeat whole blood donors from 2005 to 2014. Since few donors returned after 2500 days, we disregarded intervals exceeding this length. We derived the absolute probabilities to donate per day from the observed distribution of donation intervals (see Figure S1). These probabilities were increased x-fold and then used to obtain interval distributions for the TS and LTS donors.

Latent test-seeking donors experienced a constant rate to become activated (probability per day), but their activation was constrained to take place before day 2500. This means that the periods during which these individuals are not considering donation follow an exponential distribution. A donation interval was then randomly drawn from intervals after activation but within 2500 days.

Non-TS donors were simulated with a constant rate of infection; this rate was raised x-fold for TS donors and also for LTS donors from their day of activation onwards. The infection status of each donor was determined by his/her associated probability of being infected on the day of donation.

The sensitivity of the method to detect TS behaviour among donors was evaluated by changing various model parameters. In order to study the effects of the random nature of the donation and infection processes on the results, we performed 10 000 runs per setting (additional simulations did not impact averaged results within shown decimal precision).

As in our data analysis, the simulated intervals of infected individuals were compared to a length-biased reference distribution. The Anderson–Darling test was used to assess the likelihood of the null hypothesis being false, that is, that there is no relation between infection rate and donation interval length. This likelihood is expressed in a \( P \)-value, which is the probability that the null hypothesis is rejected by chance. The power to detect test-seeking behaviour in the donor population is the fraction of repeat runs of the model with a \( P \)-value less than 0.05 (the boundary at which we would reject the null hypothesis).

All simulations were performed in the open source software package \( R \) for statistical computing (version 3.1.2) [13]. The code used is available from the authors upon request.

**Results**

Analysing infections among Dutch donors

Of Dutch donors with a donation interval within 2005–2014, 50 acquired syphilis, 13 HIV, 4 HTLV and 2 HCV infection. The infected were more often male (80% vs. 62% of donations) but of similar age as all donors (median 48 vs. 51 years). The infected donors donated plasma rather than whole blood more often than expected: 24% of infected cases donated plasma, although only 18% of the total interdonation (at risk) time of all donors was before a plasma donation (among age and gender matched donors, two-sided exact binomial \( P = 0.18 \)). In all comparisons, we adjusted for differences in age, gender and donation type by comparing the intervals of the infected donors with the intervals of matched donors.

Of the donors who acquired an infection, relatively many had a short donation interval as compared to among all donors (Fig. 2). However, the difference between the cumulative donation interval distributions was not statistically significant according to the
Anderson–Darling test ($P = 0.54$). Therefore, we cannot conclude that any test-seeking behaviour is present among Dutch donors.

**Simulation results**

We investigated the ability of the test to detect test-seeking behaviour under different conditions. TS donors donate relatively quickly, while LTS donors return only after being exposed to an increased infection risk. When TS and LTS donors are both part of the study population, we can observe an overrepresentation of both short and long donation intervals among the infected donors (Fig. 3).

The power of the method to detect test seeking is given as the fraction of repeated simulations with a $P$-value below 0.05. We may increase this power by increasing the number of donations that are observed (see Table 1). In fact, we find that the power of our method is most sensitive to the number of infected donors observed, which means that in a population with a higher incidence rate observing fewer donations could be sufficient. For example, where at observing 150 infected donors we still miss test-seeking behaviour in one-third of cases, we would be almost certain to detect test-seeking behaviour when observing 600 infected cases.

The power to detect test seeking will obviously depend on the level of test seeking that occurs in the population as well. We recall that the proportion of donors admitting to test seeking was found to vary between 1% and 7% in questionnaire studies [1, 8]. It is also possible that while the fraction of TS donors is the same, these are engaging in even greater increased risk behaviour for infection (see Table 2). As the proportion of infected donors increases, the proportion of TS among the infected increases as well and so does the power of the test. For example, we would have been fairly certain ($P > 0.90$) to be able to identify test seeking in case at least 35 (51%) of the 69 infected donors in the Netherlands had been TS.

Our test also identifies the presence of donors that return after longer absence, who donate to obtain test results only after being exposed to increased infection risk.

---

**Table 1** Effect of the number of observed donations on the power of the AD test

<table>
<thead>
<tr>
<th>Number of donations</th>
<th>Average number infected</th>
<th>% Test seeker among infected donors</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 000 000</td>
<td>18.8</td>
<td>30</td>
<td>14</td>
</tr>
<tr>
<td>4 000 000</td>
<td>37.7</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>7 327 836</td>
<td>69.0</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>16 000 000</td>
<td>150.7</td>
<td>30</td>
<td>68</td>
</tr>
<tr>
<td>32 000 000</td>
<td>301.3</td>
<td>30</td>
<td>93</td>
</tr>
<tr>
<td>64 000 000</td>
<td>603.0</td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>

Seven percent of donations are provided by test-seeking (TS) donors, none by latent test-seeking (LTS) donors. For non-TS donors, the donation rate is based on all Dutch donors from 2005 to 2014, and the infection rate is set at 167 per 100 000 donor-years. TS donors had a twofold increased donation rate and a 10 times higher infection rate. Note that these variables were chosen such that the model would match the 69 infections in the Netherlands found among 7 327 836 donations. The power of the test is defined as the percentage of simulations with $P$-value < 0.05.
Test-seeking (TS) and latent test-seeking (LTS) donors together contribute 1 percent of all donations. Each simulation consists of 7 327 836 donations. For non-TS donors, the donation rate is based on all Dutch donors from 2005 to 2014, and the infection rate is set at 1.05 per 100 000 donor-years. Simulated TS donors had a twofold increased donation rate. Note that these variables were chosen such that the model would match the 69 infections in the Netherlands found among 7 327 836 donations, yet compared to the data presented in Table 1, a different fraction of the infected is TS. The power of the test is defined as the percentage of simulations with \( P \)-value < 0.05.

Table 3 Effect of the proportion of TS and LTS donors on the power of the AD test

<table>
<thead>
<tr>
<th>% TS</th>
<th>% LTS</th>
<th>Average number infected</th>
<th>% TS among infected donors</th>
<th>% LTS among infected donors</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>164</td>
<td>35</td>
<td>0</td>
<td>98</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>168</td>
<td>17</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>0.2</td>
<td>0.8</td>
<td>171</td>
<td>7</td>
<td>30</td>
<td>86</td>
</tr>
<tr>
<td>0.0</td>
<td>1.0</td>
<td>172</td>
<td>0</td>
<td>38</td>
<td>99</td>
</tr>
</tbody>
</table>

Naturally, mixtures of the two types of test-seeking donors will always be detectable if there are sufficient infected donors.

Test-seeking and latent test-seeking donors must differ in their donation timing in order for test seeking to be detected by our method. However, we find that the exact increase in donation rate for TS has relatively little impact on the power of the test, as long as the fraction of infected individuals who are TS remains similar (see Table S1). Similarly, the activation rate of LTS donors (in other words when inactive donors will engage in high risk, which in turn motivates them to donate) has only limited impact on the power of the test as long as activation does not occur too quickly (see Table S2). It is the fact that (L)TS donors have different donation patterns, which allows their detection in the population by our method, how much exactly their donation rates differ is of less importance.

Note that although the presence of (L)TS in a population may be revealed by inspection of donation intervals, we are not able to reveal which specific donors showed (L)TS behaviour based solely on their donation intervals. For example, with 7% TS donors in the population who have twice the daily probability to donate, the most discriminatory that we can say is that about 12% of all donors who return for donation within 60 days are TS (see Figure S2).

**Discussion**

In this paper, we describe a novel method to detect test-seeking behaviour in a population of repeat blood donors by analysing donation interval data that by most blood banks are collected on a routinely basis. If identified by our method, test-seeking behaviour would ideally be verified, for example in interviews of infected donors [9, 10]. Yet the great advantage of our method is that it does not rely on the willingness of individuals to disclose such socially denounced behaviour. One limitation of the method is that to detect test-seeking behaviour, a substantial number of infections must be observed. Therefore, this method may be a useful addition especially in settings where infection rates are relatively high.

Because of the limited number of infections observed in the Netherlands, we pooled all infection types in order to achieve the greatest obtainable power. Although most donors were infected with syphilis, test seekers among these may have worried about a potential HIV or other infection, as the same sexual risk behaviour (e.g. unprotected sex with an incidental partner) may lead to syphilis, HIV, HTLV and Hepatitis C or B virus infection [10, 14]. Hepatitis and HIV are also transmitted by drug injection [15], and drug using test seekers may have different
donation patterns compared with those with increased sexual risk behaviour. This does not detract from the validity of our method, as we test for any association between the timing of donation and the risk of being infected. If some of the infections are less strong or even unrelated to test-seeking behaviour, this would merely dilute the signal and therefore reduce somewhat the ability to detect the presence of this effect in the data.

Despite having combined all infections in our analysis, we have to conclude that there is no statistical evidence for test-seeking behaviour among Dutch donors. Indeed, the analysis seems to confirm that test seeking is at least not very common in the Netherlands, since we would have expected to be able to show test seeking if most of the 69 infected donors had been TS. A lower extent of test-seeking behaviour would less likely have been detected by our test.

The assessment of the Dutch donor population showed a trend that would fit test seeking more than latent test seeking. However, this might be due to the limited time period considered. In an additional analysis, where we added donors that were infected in the years 2005–2014 but who had their last negative donation before 2005, we found a relatively large number of long intervals (>2 years) among the infected. It should be noted that this analysis may be biased by changes in the general donation behaviour over time (see the Supporting information). Where specifically TS or rather LTS are expected, performing the one-sided Kolmogorov–Smirnov test gives somewhat more power than the Anderson–Darling test for detecting deviations from regular donation patterns [16].

The interpretation of the results must be done with care as other effects might also distort the intervals of infected donors relative to all donors. Holidays outside of Europe, for example, lead to temporary donation deferral and such donors may also have increased infection risks. This could lead to changes in the distribution of donation intervals similar to changes as a result of the presence of (short term inactive) LTS donors. Such patterns, however, were not identified in the donation sequences of infected Dutch donors.

Other ways of looking at donation intervals can also be valuable. Where we compared donation intervals at the population level, Schreiber et al. [17] compared the intervals within an individual’s donation career. They found that for HIV-infected donors the interval during which infection occurred was longer than expected from their own previous (infection free) intervals. They speculated this finding to be due to the acute HIV syndrome (flu like symptoms) leading to postponement of donation.

Using the method by Schreiber et al., we found that of the infected Dutch donors 52% (27 of 52 infected donors with multiple returns, exact binomial $P = 0.45$) returned later than expected based on their own previous intervals (for HIV-infected donors, this proportion was 64%, $P = 0.27$). This seems surprising given the quicker return of these infected donors when compared to all donors (see Fig. 2). An explanation could be that such donors are consistently TS (thus over multiple donations). This would explain the relatively short donation intervals of those infected when compared to the general donor population, despite relatively long donation intervals of infected donors when compared to their own donation histories.

The fact that a relatively high number of infected donors donated plasma rather than whole blood (though not statistically significant) might also indicate the presence of such persistently high-risk donors who aim at being tested on a regular basis.

Our method has to be adapted before it can be applied to infections that can be cleared and are then no longer detected. When presuming that all donors that ever obtained HBV infection tested positive (i.e. ignoring the effect of donors clearing this infection), we observed a substantial increase in short donation intervals [18, 19]. This is because most infections that occurred during longer intervals had cleared by the time the donor provided a next donation. One can still test for the presence of test-seeking behaviour in the donor population when correcting the reference distribution for the impact of clearing, or when screening for anti-HBcore (see Supporting information).

The method described in this paper does not allow identification of individual test seekers, but only allows assessment of the extent of test-seeking behaviour within a donor population. In case such behaviour is detected, extra measures may be implemented to identify and/or exclude test-seeking donors. These might include more careful donor screening (questionnaires and interviews) [3], donor education [20] or perhaps offering free (anonymous) testing to non-donors [21]. Measures may be tailored to TS or LTS behaviour observed among infected donors (e.g. stricter screening of donors upon returning from a longer period of absence from donation).

The seeming lack of test seeking in the Netherlands may be due to the adequacy of current donor selection procedures and the availability of testing through regular public healthcare facilities in our country. It would be very interesting to apply this method to data from other regions, where donor screening may be less adequate [6, 7].

In conclusion, the proposed method, which requires only an easy-to-perform assessment of routinely collected data, provides an additional tool that enables identification of risk behaviour among donors that is difficult to assess in any other way, making the invisible visible.

**Conflicts of interest**

The authors declare no conflict of interests.
References

21. Gonzalez TT, Blatyta PF, Santos FM, et al.: Does offering human immunodeficiency virus testing at the time of blood donation reduce transfusion transmission risk and increase disclosure counseling? Results of a randomized controlled trial, São Paulo, Brazil. Transfusion 2015; 55:1214–1222

Supporting Information

Additional Supporting Information may be found in the online version of this article:
Table S1: Effect of the TS donation rate multiplication on the power of the AD test.
Table S2: Effect of the LTS activation rate on the power of the AD test.
Fig. S1. The hazard rate to make a subsequent whole blood donation for all Dutch donors who donated within 2005–2014.
Fig. S2. The probability to be a test-seeking donor based on donating within a certain number of days from the last donation, and on infection status at donation.
Fig. S3. The cumulative distribution of last donation intervals of HBV infected donors compared to the cumulative reference distribution from the inter-donation intervals of all Dutch donors in 2005–2014.
Fig. S4. The cumulative distribution of last donation intervals of infected donors tested positive in 2005–2014 (including those with a pre-seroconversion donation before 2005) compared to the cumulative reference distribution from the inter-donation intervals of all Dutch donors returning in 2013–2014.

© 2016 International Society of Blood Transfusion
Var Sanguinis (2016) 111, 274–280