Transfusion-related acute lung injury : etiological research and its methodological challenges

Middelburg, R.A.

Chapter 4

A solution to the problem of studying blood donor related risk factors when patients have received multiple transfusions

Rutger A. Middelburg
Saskia le Cessie
Ernest Briët
Jan P. Vandenbroucke
Johanna G. van der Bom

Transfusion 2010; 50: 1959-66

Abstract

Background

A problem when studying adverse events of blood transfusions is that patients have usually received transfusions from several donors while only one of these donors is the actual cause. This will result in underestimation of the effect of donor related risk factors if not adequately corrected for. We encountered this problem when studying transfusion-related acute lung injury (TRALI) and describe four methods to overcome this problem.

Study design and methods

Simulated data are used to illustrate the results of six different approaches: not correcting for the number of donors, using standard correction methods, and four newly proposed methods. Donor sex is used throughout as an example. The first two new methods apply restriction of the study to cases who have received a transfusion from a single donor or from donors who are all of the same sex. In both restriction designs the sex of the causal donors is known and can be compared to the expected value from a reference population. The other two new methods apply statistical correction for the number of donors, either by standardization or by maximum likelihood methods.

Results

If not corrected for, or if corrected for by standard methods, increasing numbers of donors per patient result in decreasing estimates of the effect of risk factors. All four newly proposed methods yield valid estimates.

Conclusion

It is clear that the problem of multiple transfusions requires specialized correction methods. All four newly proposed methods yield on average good estimates of the underlying true value.

Introduction

Many adverse events associated with blood transfusions are due to one single transfusion, like transfusion transmitted infections, allo-immunization, most anaphylactic and allergic reactions, and transfusion-related acute lung injury (TRALI). The majority of transfusion recipients receive more than one transfusion. In most cases of adverse events it is not possible to identify the single causal transfusion. This complicates many studies that examine the association between donor related risk factors and transfusion reactions. We encountered this problem when studying TRALI.

TRALI is a form of acute respiratory distress syndrome (ARDS) that develops during, or within six hours after, the transfusion of one or more blood products.[1] TRALI is currently recognized as the most common of the severe side effects of blood transfusion.[2-5] It has an estimated incidence of 1 in 5,000 transfusions and a mortality commonly estimated to be between 5 and 10%, with the majority of patients spontaneously recovering within 96 hours, without long term sequelae.[1,6,7]

Both patient and donor related risk factors are thought to be involved in TRALI. Of the donor related risk factors, donor derived leukocyte antibodies are thought to be the most important.[1,8,9] Leukocyte antibodies are almost exclusively found in donors who have previously been exposed to allo-antigens and the large majority of allo-exposed donors are women who have been pregnant.[10,11] Therefore, donor sex is considered an important donor related risk factor for TRALI and several countries have excluded female donors from donation of plasma for transfusion. Since it is so important and can be determined so easily, donor sex will be used as an example throughout this paper.

Obtaining a quantitative estimate of the contribution of female donors to the occurrence of TRALI is complicated by the fact that in most cases TRALI patients have received transfusions from more than one donor before developing TRALI. Figure 1 gives a histogram of the number of donors involved in each case of a Dutch case series of 86 TRALI patients. In this representative case series 85% of TRALI patients have received transfusions of more than one donor in the six hours before the onset of symptoms. All donors of blood products transfused in this six hour window have to be considered as potentially causal. However, due to the low incidence of TRALI it can be assumed that the probability of having two causal transfusions is negligibly small. It is generally accepted that TRALI is caused by a transfusion from only one of the donors, the causal donor. Consequently, without identification of this causal donor, the crude quantitative estimate of the contribution of female donors to the occurrence of TRALI will be an underestimation.

The dilution effect resulting from transfusions from multiple donors can be quite substantial and conventional methods to correct for the number of transfusions (like stratifying by the number of transfusions or adding the number of transfusions in a

regression model) do not result in adequate correction. We discuss four approaches to obtain valid estimates of the contribution of female donors to the occurrence of TRALI. We simulated 1,000 data sets, each comprising 1,000,000 transfusion recipients, to describe the dilution effect of multiple transfusions, to illustrate the inadequacy of standard correction methods, and to compare the performance of the four newly proposed methods.

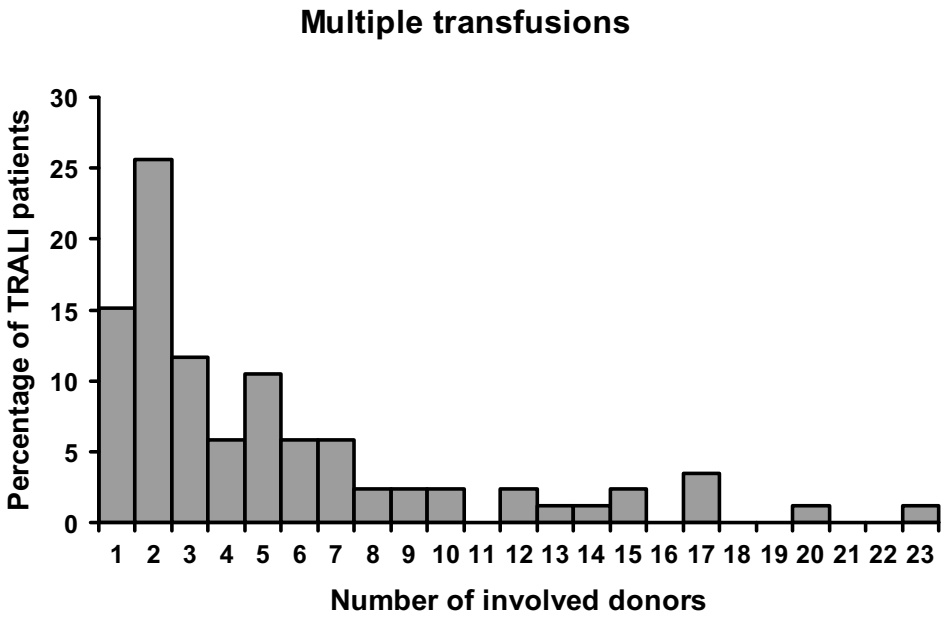

Figure 1: The number of donors involved in a series of 86 Dutch TRALI patients.

Materials and Methods

We will present here an informal description of the comparisons made between donors associated with TRALI cases and the source population of these donors. For a mathematical presentation of these comparisons we refer the interested reader to the Appendix. In the text we will also present simple numerical examples with each method. For the first three methods the actual calculations can be demonstrated in these examples; the last method (Correction by maximum likelihood estimation) requires specialized software to maximize the likelihood formula and the actual calculations can therefore not be presented. Thereafter, we will present the large sample simulations.

Assumptions and definitions

Two types of TRALI patients

We assume that TRALI can be caused by a women specific mechanism (i.e. antibodies as a consequence of pregnancies) or by unspecified other mechanisms (i.e. either antibodies due to other types of immunizing events, or other causes all together). Thus some TRALI cases are caused by female donors and could have been prevented by the exclusion of female donors. Such cases are caused by a women specific mechanism and will be referred to as "type I TRALI" cases. All other TRALI cases will be referred to as "type II TRALI". Amongst the type II TRALI cases there are also TRALI patients who have received blood from one or more female donors. However, in these cases the sex of the donor was a coincidence, rather than a causal prerequisite (i.e. either none of the female donors was causal, or one of the female donors was causal, but not due to a women specific mechanism). These type II TRALI cases were, therefore, caused independent of donor sex.

Problem of multiple donors

In standard etiological studies the exposure prevalence among cases is compared with that of the source population of the cases in a twofold table (table 1). All figures in table 1 represent numbers of patients, either with or without disease and with or without exposure to a female donor. If a TRALI patient (in the table referred to as TRALI +) received blood from only one donor and that donor was female, the patient is considered a female exposed case (A) and if the only donor was male the patient will be considered a female unexposed case (B). However, if the TRALI patient received more than one transfusion the causal transfusion is not known. Therefore, the values for cells A and B are not known and table 1 can not be composed directly. The four methods we propose describe different methods of dealing with this problem.

Table 1: Patients with TRALI, who have received transfusions from a single donor (TRALI +) and reference patients who also received a transfusion (TRALI -) according to the sex of one donor

TRALI	Female donor	Male donor	Total
+	A	B	N_1
-	C	D	N_2
Total	M_1	M_2	T

All figures represent numbers of patients.

Data assumed available or computable

We assume that for each TRALI patient with multiple donors we know the numbers of female and male donors. These observed numbers of female and male donors will need to be compared with expected numbers of female and male donors. The expected numbers can be estimated from the number of transfusions received and the fractions of female and male donations, as can be documented from the relevant donor population. Relevant in this context means: the donor population that represents the expected sex distribution for the donors involved in the TRALI case. The sex distribution of donors can be different between different countries or regions and between different product types and can also change over time. Therefore, the relevant donor population will, in practice, be a donor population donating the product type received by the TRALI patient, in the same geographical area, and during the same period as the TRALI occurred.

Definition of measure of effect

Our aim is to estimate the population attributable risk (PAR, the fraction of type I TRALI) as a measure of the contribution of female donors to the occurrence of TRALI. Estimation of the population attributable risk in the first two methods (restriction based methods) requires the relative risk (RR) to be estimated first. In the last two methods (correction based methods) the population attributable risk is estimated directly and the RR is calculated only for purposes of comparing the four methods' performance.

Proposed methods

1. Restriction to single donor cases

The first and simplest method to remove the diluting effect of multiple donors is to restrict the study to TRALI cases who have received transfusions from only one donor in the six hour period preceding the onset of symptoms. Since this one donor is by definition the causal one, the sex of this donor can be compared directly with the expected fraction of female donors. The expected value equals the fraction of blood products donated by female donors in the reference population. If the majority of TRALI cases would be caused by unspecified mechanisms (i.e. non-women specific mechanisms), the fraction of female donors among cases will be similar to that in the reference population and the relative risk (RR) will be close to unity.

Since the problem of multiple transfusions is effectively removed from the data, the population attributable risk can be calculated using standard formulas (Appendix, equations 1 and 2). The standard error and confidence interval for the estimated population attributable risk can be calculated using the delta method.[12]

For example, if the fraction of blood products donated by female donors is 0.35 and 20 patients have each received only a single transfusion before developing TRALI, 8 of which were from female donors. The RR would be 1.2 (i.e. [8x0.65]/[12x0.35]) and the population attributable risk would consequently be 8% (i.e. [0.35x1.2-0.35]/[0.35x1.2+0.65]).

2. Restriction to unisex cases

The second method also involves restriction of the study population to selected TRALI cases. In this case TRALI patients who have received transfusions either only from male donors or only from female donors (unisex cases) are included and compared to the reference population. If all donors involved in a TRALI case are of the same sex, this sex must be the sex of the causal donor. Therefore, the sex of the causal donor is known, even if the causal donor is not explicitly identified.

Since the sex of the causal donor is known, the relative risk of TRALI for a transfusion from a female donor versus a male donor can be calculated. However, the probabilities of receiving all transfusions either only from male donors or only from female donors are not necessarily equal to the fractions of donations made by female and male donors. Instead we should determine the relative probabilities of receiving multiple transfusions from only female and only male donors, given that we already know this TRALI patient to be a unisex case. These probabilities can easily be calculated if we know the number of transfusions received and the fractions of donations made by female and male donors (Appendix, equations 3 and 4). Further calculations of the RR and population attributable risk are identical to those used after restriction to single donor cases.

Consider the next example; if the fraction of blood products donated by female donors is again 0.35 and 20 patients have each received three transfusions, either only from male donors or only from female donors, before developing TRALI. Of these 20 patients 8 received all three transfusions from female donors. Since we have selected unisex cases we have to compare the number of cases caused by female and male donors to the number of unisex transfusion recipients, without TRALI, who received the same number of transfusions from female and male donors (Appendix, equations 3 and 4). Out of 20 unisex recipients receiving three transfusions each we would expect 2.7 to have received all three transfusions from female donors (i.e. $0.35^3/[0.35^3+0.65^3]$). The RR would be 4.3 (i.e. [8x17.3]/[12x2.7]) and the population attributable risk would consequently be 31% (i.e. [2.7x4.3-2.7]/[2.7x4.3+17.3]).

3. Correction by standardization

The third method applies *direct standardization*, according to the exposure in the reference population, to correct for the number of transfusions received by each TRALI patient. It consists of three steps. First, we estimate the total number of TRALI patients for whom the

causal transfusion was from a female donor (Appendix, equation 5). This number comprises all patients with a type I TRALI, i.e. those in which case the causal transfusion by definition has to be from a female donor, and a number of type II TRALI cases, i.e. in whom the causal transfusion is from a female donor by chance.

Second, from the fractions of female donors in the different reference populations (which can be different for each TRALI patient), we calculate a weighted average (Appendix, equation 6). This is the fraction of cases that is expected to be exposed. Multiplying this fraction by the number of cases will give the expected number of exposed cases, needed for the standardization.

Thus we have estimated both the observed number of exposed cases and the expected number of exposed cases. The third and final step is to use these numbers to calculate the standardized population attributable risk and RR (Appendix, equations 7 and 8)

To derive confidence intervals for the population attributable risk, the variance of the population attributable risk can be estimated as shown in the Appendix (equations 9-11). Another option would be to apply bootstrapping procedures. Furthermore, when either the fraction of female donors or the contributions of each stratum to the population attributable risk are reasonably homogeneous across strata, the normal non-weighted average of the fraction of female donors can be used directly, without the need to calculate a weighted average. Homogeneity of the contributions to the population attributable risk across strata is also referred to as absence of effect modification. Effect modification and its impact on the choice of method to use for correction will be considered in more detail in the discussion.

Consider again the example of 20 TRALI patients, each receiving three transfusions, while the fraction of blood products donated by female donors is 0.35. From all 60 involved donors 24 are again female, but we now assume these are randomly distributed across all patients. Therefore, four patients have received only transfusions from male donors, nine have received a single transfusion from a female donor, six have received two transfusions from female donors and a single TRALI patient has received all three transfusions from female donors. Since most patients have received transfusions from both female and male donors, it is impossible to tell which of these TRALI cases was caused by male or female donors. However, we can estimate that in 10 of our 20 TRALI cases the causal donor must have been female (Appendix, equation 5: 4x[0-3x0.35+0.35]+9x[1-3x0.35+0.35]+6x[2-3x0.35+0.35]+1x[1-3x0.35+0.35]). Note, though, that in 7 of these 10 cases, although the causal donor was female, the sex of the donor was a coincidence and the TRALI was not caused by a women specific mechanism. Since the fraction of blood products donated by female donors was constant for all cases the average (weighted or non-weighted) also equals 0.35. The population attributable risk is thus 23% (i.e. [10-20x0.35]/[20-20x0.35]) and the RR is 1.9 (i.e. [0.23/[0.77x0.35]]+1).

4. Correction by maximum likelihood estimation

The fourth method involves the use of a statistical model and maximum likelihood methods to correct for the number of transfusions received by each TRALI patient. The only data needed for this model is the sex of all the donors involved in each TRALI patient.

In this model there are two unknown parameters, which are estimated simultaneously. Firstly, the fraction of female donors in the reference population. Since this fraction is estimated from the available data on the donors of TRALI patients, this method does not require a reference population to be defined. Secondly, the population attributable risk is the fraction of TRALI cases preventable by the exclusion of all female donors.

For a patient with type I TRALI the causal donor is female. The number of female donors among the remaining donors of this patient follows a binomial distribution. For a patient with type II TRALI, the total number of female donors will follow a binomial distribution. The probability that a given TRALI is of type I is the population attributable risk (PAR) and the probability that it is of type II is (1-PAR). Together we can use this information to compose a likelihood formula for the probability of observing a given number of female donors out of a total number of donors for a person for whom the type of TRALI is unknown (Appendix, equations 12-15).

Maximizing this function with respect to the population attributable risk and the fraction of female donors in the reference population, yields the maximum likelihood estimates for both. Numerical methods are needed to maximize this function; we used the function "optim" of the statistical package R.[13] The second derivatives of the log-likelihood function can be used to estimate the standard errors of the estimated parameters, however in this particular case it is easier to calculate profile-likelihood based confidence intervals.[14]

Maximum likelihood methods can also be used when the expected fraction of female donors is obtained from population data, as was also the assumption for the other three methods. In this case the likelihood function has only one unknown parameter (i.e. the population attributable risk).

We consider again the previous example of 20 TRALI patients, each receiving three transfusions, while the fraction of blood products donated by female donors is 0.35. From all 60 involved donors 24 are again female and again four patients have received only transfusions from male donors, nine have received a single transfusion from a female donor, six have received two transfusions from female donors and a single TRALI patient has received all three transfusions from female donors. When we applied the standardization method (previous method), the result was a population attributable risk of 23%. If we determine the value of the population attributable risk that gives the maximum result from the likelihood formula (with a fraction female donations of 0.35) this would be almost identical at 24%.

Data simulation

We designed a simulation study in which we created 1,000 datasets with 1,000,000 patients each. We used the statistical package R to produce these data sets.[13]

1,000,000 patients were set to receive a random number of transfusions in the range from one to ten, using a uniform distribution. The fraction of donations by female donors in the total population of donors was set at 0.35, based on the Dutch donor population. The number of female donors amongst the number of received transfusions was subsequently drawn from a binomial distribution.

The overall probability of developing TRALI was set at 1/5,000 transfusions[1] and the relative risk associated with a transfusion from a female donor was arbitrarily set at 10. This implies that the probability of developing TRALI after a transfusion from a male donor equals 1/20,750 transfusions ([1/5,000]/[10x0.35+0.65]). The chance of developing TRALI after a transfusion from a female donor is consequently 10/20,750 transfusions. The true population attributable risk can be obtained using equation 1 and equals 75.90 % (i.e. [3.5-0.35]/[3.5+0.65]=3.15/4.15).

We had two primary effect measures of interest in our study: the fraction of TRALI cases preventable by excluding all female donors (population attributable risk, PAR) and the relative risk of TRALI after a transfusion from a female donor compared to a transfusion from a male donor. We applied all four proposed approaches to all 1,000 data sets. For each of the resulting eight different estimates the median and 5th and 95th percentiles of the estimates of the 1,000 simulations were determined. The median of 1,000 estimates gives an impression of the bias of that approach and the interval between the 5th and 95th percentiles about the precision.

R codes for running the four methods can be obtained from the corresponding author.

Results

True values

In our simulated data sets the true relative risk (RR) was 10 and the true population attributable risk (PAR, i.e. the fraction of TRALI cases preventable by the exclusion of female donors) was 75.90%.

Crude estimate and stratification by number of transfusions

Analyzing the simulated data without correction for the number of transfusions resulted in a crude RR of 1.3 (5th to 95th percentile: 1.3 to 1.4) and a crude population attributable risk of 10.87% (9.48% to 12.22%) (table 2).

When the RR was calculated within strata of the number of transfusions it decreased exponentially with each additional transfusion (figure 2). A conventional method of correction for the number of transfusions consists of pooling these stratum specific RR. Pooling resulted in an overall RR of 1.5 (1.4 to 1.5) and a population attributable risk of 13.82% (12.22% to 15.27%) (table 2), which was hardly better than the unadjusted RR and population attributable risk.

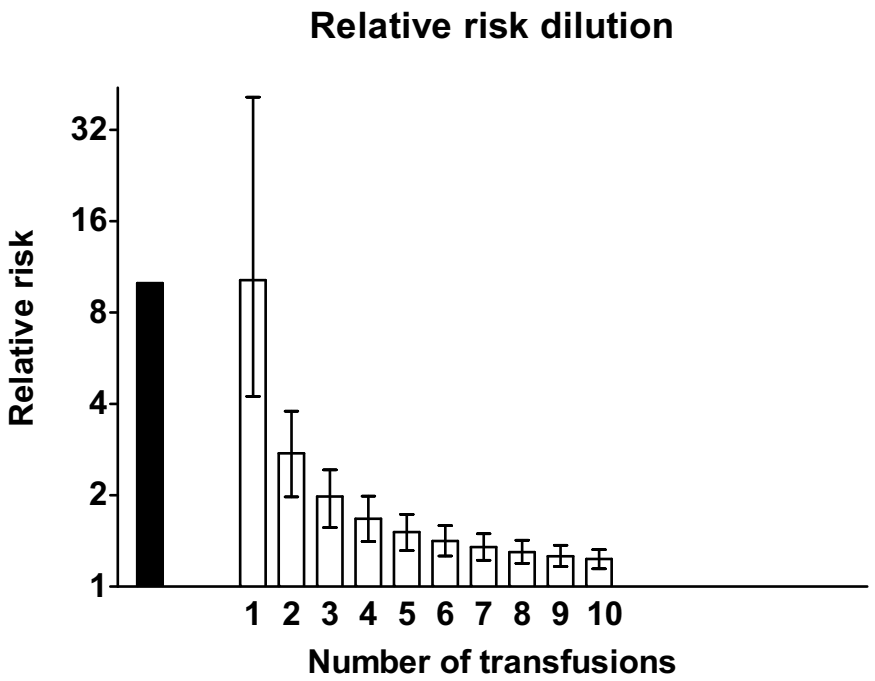

Figure 2: Dilution of the relative risk of female donors on TRALI with increasing number of transfusions, compared to the true value. The true value is represented by the filled bar. Error bars represent 5th and 95th percentiles of estimates from 1,000 simulations.

Proposed methods

Applying the four suggested methods to the simulated data resulted in good estimates of the population attributable risk and the RR (table 2). The main difference between the results was the width of the interval between the 5th and 95th percentiles of 1,000 simulations (table 2).

A median total of 1,100 TRALI cases (5th to 95th percentile: 1,044 to 1,150) was created per simulation. Of these cases a median of 20 cases (13 to 27) were single donor

cases and a median of 64 cases (51 to 77) were unisex cases. All single donor cases are inevitably also unisex cases. The two restriction based analyses were based on these cases. In the standardization and maximum likelihood analyses all cases could be included, resulting in more precise estimates of the RR and population attributable risk (table 2).

In the maximum likelihood analyses it was also possible to estimate the fraction of female donations in the reference population simultaneously with the population attributable risk. The median of the estimates of this fraction was 0.3496 (5^{th} to 95^{th} percentile: 0.3372 to 0.3634), while the true fraction of female donations in the simulated data was 0.3500 (0.3497 to 0.3504). Using the true fraction instead of estimating it simultaneously with the population attributable risk resulted in a further increase of the precision of the estimated population attributable risk, but had no material effect on the bias of the population attributable risk (table 2).

Table 2: Results of 1,000 simulations to compare the performance of the four different suggested approaches with the results after conventional correction and with the unadjusted result

Method of analysis	RR (p5-p95)*	PAR (p5-p95)*
True value	10	75.90
Unadjusted	1.3 (1.3-1.4)	10.87 (9.48-12.22)
Conventional adjustment †	1.5 (1.4-1.5)	13.82 (12.22-15.27)
Single donor	10.23 (4.2-41)	75.72 (52.00-91.90)
Unisex	10.13 (6.3-18)	76.16 (65.04-85.79)
Standardization	10.12 (6.7-17)	76.14 (66.42-85.15)
Maximum Likelihood	10.12 (7.5-15)	76.15 (69.32-82.60)
ML, *p* unknown ‡	10.07 (6.6-16)	76.15 (66.83-84.12)

* (p5-p95) represent 5^{th} and 95^{th} percentiles of estimates from 1,000 simulations.

† Conventional correction method using stratification according to the number of transfusions and subsequent pooling.

‡ Maximum likelihood method assuming the fraction of female donations (*p*) from the reference population to be unknown, therefore estimating the population attributable risk and *p* simultaneously.

RR: Relative risk. PAR: Population attributable risk.

Discussion

From the crude results obtained without correction it is clear that the problem of multiple transfusions can not be ignored when studying donor related risk factors for TRALI. Further, it is shown that conventional methods of correction for the number of transfusions do not solve the problem. All four proposed methods for dealing with this problem yield excellent estimates of the true fraction of TRALI cases preventable by the exclusion of female donors.

Conventional methods

Conventional methods (like stratifying by the number of transfusions or adding the number of transfusions in a regression model) take the weighted average of the relative risks shown in figure 2. Therefore, even if only a few patients have received more than one transfusion, the relative risk "corrected for the number of transfusions" by conventional methods will already be biased. If there are any patients at all who have received more than a single transfusion, one of the four proposed methods should always be used.

Restriction methods

The restriction methods are the most intuitively clear and computationally easy but, by definition, put a further restraint on the already limited number of TRALI cases available. Single donor cases are not only rare but any TRALI caused by the transfusion of pooled products (i.e. in some blood services either platelets, plasma, or both) can not be studied in this way.

Unisex cases occur more often, which likely contributed to the fact that this approach has been applied to real data, in a study of internationally gathered unisex cases.(*submitted*) However, for other donor related risk factors (i.e. parity) information is not available for all donors. This information has to be gathered specifically for all donors involved in TRALI cases to be able to identify cases receiving either all transfusions from parous donors, or all transfusions from nulli-parous donors. When known, this information can be used more efficiently by applying one of the other methods for correction.

Standardization

If information on the reference group is available, standardization can be used as a method for correction. This method uses the data a bit less efficiently than maximum likelihood estimation, but offers the advantage of relatively easy and straightforward calculations. Moreover, as with all standardization methods, it is the only valid summary measure in the case of effect modification. For instance if the risk associated with female donors of plasma rich products is different from the risk associated with female donors of red

cells.(*submitted*) In this case it will provide a summary measure which is the weighted average of stratum specific measures, weighted for the actual composition of the total donor population which was used for standardization. It is this measure that most accurately tells a blood bank which effect to expect from deferral of all female donors. However, it is only a valid estimate in the population it was determined in, or a population with very similar distribution of female donors across product types and very similar product type usage. In contrast, the results of the other three methods are often considered etiologically more relevant, since the RR estimated from these methods is less dependent of the population in which it was determined. However, since the population attributable risk is always dependent on the exposure prevalence in the population it does not share the RR's advantage of greater validity in other populations. The advantage belongs therefore to the RR only, while the population attributable risk is likely to be of greater interest to blood banks. However, the standardization method also has one drawback, common of all standardization methods. The reference population needs to be large enough to avoid estimates of the exposure fraction to equal zero or one hundred percent in even a single stratum. This is generally not a problem when studying donor related risk factors.

Maximum likelihood estimation

Maximum likelihood estimation uses the data most efficiently, resulting in the most precise estimate of the effect. Another advantage of the maximum likelihood method is that it doesn't necessarily require information on the reference population. The fraction of female donors in the reference population and the population attributable risk can be estimated simultaneous from information on donors involved in TRALI cases alone. Although estimation of two variables does decrease the precision of the estimate, this could be a worthwhile tradeoff in some situations. Especially when investigating other donor related risk factors, such as leukocyte antibodies. On such risk factors information might already be gathered for donors involved in TRALI cases, but could be expensive or cumbersome to collect for a large reference group.

Conclusions

In conclusion, all four methods can be used to study the contribution of donor related risk factors to the occurrence of TRALI. The unisex method is reliable and computationally easy, but can be difficult to apply to other donor related risk factors. Furthermore, since pooling across strata with different relative risks is only allowed with the standardization method, this will often be the only method to provide one summary measure. In these instances the standardization method provides the most relevant estimate of the effect on TRALI incidence that can be expected from implementing new donor deferral policies,

especially if these policies will be implemented on the same population used in the standardization.

Acknowledgements

The authors thank two anonymous reviewers for invaluable comments and suggestions on this paper.

References

1. Popovsky MA, Moore SB. Diagnostic and pathogenetic considerations in transfusion-related acute lung injury. Transfusion 1985 Nov;25(6):573-7.
2. SHOT. Serious Hazards of Transfusion (SHOT): SHOT Annual Report 2008 2009 [cited 2010 Jan 19] Available from http://www.shotuk.org/home.htm.
3. TRIP. "Transfusie Reacties in Patiënten" (TRIP): TRIP rapport 2008 2010 [cited 2010 Jan 19] Available from http://www.tripnet.nl/.
4. Engelfriet CP. Haemovigilance. Vox Sang. 2006 Apr;90(3):207-41.
5. Eder AF, Herron R, Strupp A, Dy B, Notari EP, Chambers LA, Dodd RY, Benjamin RJ. Transfusion-related acute lung injury surveillance (2003-2005) and the potential impact of the selective use of plasma from male donors in the American Red Cross. Transfusion. 2007 Apr;47(4):599-607.
6. Moore SB. Transfusion-related acute lung injury (TRALI): clinical presentation, treatment, and prognosis. Crit Care Med. 2006 May;34(5 Suppl):S114-S117.
7. Wallis JP. Transfusion-related acute lung injury (TRALI): presentation, epidemiology and treatment. Intensive Care Med. 2007 Jun;33 Suppl 1:S12-S16.
8. Bux J. Transfusion-related acute lung injury (TRALI): a serious adverse event of blood transfusion. Vox Sang. 2005 Jul;89(1):1-10.
9. Silliman CC, Boshkov LK, Mehdizadehkashi Z, Elzi DJ, Dickey WO, Podlosky L, Clarke G, Ambruso DR. Transfusion-related acute lung injury: epidemiology and a prospective analysis of etiologic factors. Blood 2003 Jan 15;101(2):454-62.
10. Boulton-Jones R, Norris A, O'Sullivan A, Comrie A, Forgan M, Rawlinson PS, Clark P. The impact of screening a platelet donor panel for human leucocyte antigen antibodies to reduce the risk of transfusion-related acute lung injury. Transfus.Med. 2003 Jun;13(3):169-70.
11. Densmore TL, Goodnough LT, Ali S, Dynis M, Chaplin H. Prevalence of HLA sensitization in female apheresis donors. Transfusion 1999 Jan;39(1):103-6.
12. Rothman KJ, Greenland S. Attributable fraction estimation. In Modern Epidemiology. 2 ed. 1998. p. 295-7.
13. R: A language and environment for statistical computing R Development Core Team (2007). Version 2.6.1. 2007 Nov 26.
14. Venzon DJ, Moolgavkar AH. A method for computing profile-likelihood based confidence intervals. Applied Statistics 1988;37:87-94.
15. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J.Natl.Cancer Inst. 1959 Apr;22(4):719-48.
16. Miettinen OS, Cook EF. Confounding: essence and detection. Am.J.Epidemiol. 1981 Oct;114(4):593-603.

Appendix: Mathematical presentation of the four proposed methods

Data assumed available or computable

In standard etiological studies the exposure prevalence among cases is compared with that of the source population of the cases in a twofold table (table 1). Appendix table 1 instead represents these data for a single patient in a twofold table relating the sex of donors to the occurrence of TRALI. In this table *a* and *b* represent observed numbers of female and male donors and *c* and *d* are expected numbers of female and male donors for a patient with an equal total number of transfusions. These expected values are estimated from the fractions of female and male donations from the relevant donor population and the number of transfusions. In fact *c* and *d* can also be interpreted directly as expected fractions of female and male donors, instead of as expected numbers. For convenience this interpretation will be used throughout since it allows for consistent and short notations in all further definitions and equations.

For an analysis with N strata (i.e. N TRALI cases) the stratum specific values for the i^{th} stratum (i=1, i=2, ..., i=N) are given by:

a_i = the number of female donors (Appendix table 1).

b_i = the number of male donors (Appendix table 1).

n_{1i} = the total number of donors (a_i+b_i; Appendix table 1).

p_i = the fraction of female donors in the reference population.

From this, the following can be calculated:

c_i = the expected fraction of female donors for a given TRALI patient.

d_i = the expected fraction of male donors for a given TRALI patient.

Furthermore:

A_i = the (not always directly observed) contribution that this case will make to cell A of table 1 of the main paper.

Definition of effect estimator

Our aim is to estimate the population attributable risk (PAR, the fraction of type I TRALI) as a measure of the contribution of female donors to the occurrence of TRALI. It can be shown[12] in a population with a fraction *p* of female donors that the PAR equals:

$$\text{PAR} = \frac{pRR - p}{pRR + (1 - p)} \qquad (1.)$$

With RR the relative risk of developing a TRALI after a transfusion from a female donor, compared to a transfusion from a male donor.

Appendix table 1: Numbers of donors involved in one case of TRALI (TRALI +) and reference donors from the relevant donor population (TRALI -) according to sex

TRALI	Female donor	Male donor	Total
+	a	b	n_1
-	c	d	n_2
Total	m_1	m_2	t

All figures represent numbers of donors, but c and d can also be interpreted directly as the fractions of female and male donors in the relevant donor population. In this case n_2, m_1, and m_2 also loose their interpretation as numbers of donors. Since n_2, m_1, and m_2 are not necessary for any of the subsequent calculations, this distinction can be ignored.
Each TRALI patient has his or her own table. For an analysis with N strata (i.e. N TRALI cases) the stratum specific values for the i^{th} stratum (i=1, i=2, ..., i=N) are given by a_i, b_i, n_{1i}, c_i, and d_i.

Restriction to single donor cases

In this situation, the RR needed for equation 1 can be directly estimated using a Mantel Haenszel estimate (RR_{MH})[15] calculated, as given by:

$$RR_{MH} = \frac{\sum_{i=1}^{N} A_i d_i}{\sum_{i=1}^{N} B_i c_i} \qquad (2.)$$

With a_i=A_i, b_i=B_i, p_i=c_i and d_i=1-c_i. By estimating p by $p = \Sigma p_i/\Sigma n_{1i}$ (the average of p_i) across all strata the population attributable risk of female donors can be calculated by completing equation 1.

Restriction to unisex cases

In this situation, we need to estimate the probability of receiving all transfusions either only from male donors or only from female donors. This means we should determine the probabilities of receiving transfusions from female or male donors only, given that we already know this TRALI patient to be a unisex case. We know p_i to the power n_{1i} gives the probability of receiving all n_{1i} transfusions from female donors and (1-p_i) to the power n_{1i} gives the probability of receiving all n_{1i} transfusions from male donors. Adding these probabilities gives the probability of being a unisex case (with either male or female donors). Dividing either of the previous two probabilities by the probability of being a unisex case, gives the probability of receiving all transfusions from female or male donors, conditional on being a unisex case:

$$c_i = \frac{p_i^{n_{1i}}}{p_i^{n_{1i}} + (1-p_i)^{n_{1i}}} \quad (3.)$$

$$d_i = \frac{(1-p_i)^{n_{1i}}}{p_i^{n_{1i}} + (1-p_i)^{n_{1i}}} = 1\text{-}c_i \quad (4.)$$

Since all donors are of the same sex and only one donor per stratum is considered causal A_i and B_i can be estimated by: $a_i/n_{1i}=A_i$ and $b_i/n_{1i}=B_i$.

Correction by standardization

The third method applies direct standardization, according to the exposure in the reference population, to correct for the number of transfusions received by each TRALI patient. It consists of three steps. First, we estimate A (equation 5): the number of exposed TRALI patients in table 1. Second, we calculate the weighted average of p_i (p^*: equation 6). This is the fraction of cases that is expected to be exposed. Multiplying this fraction by the number of cases (N_1) will give the expected number of exposed cases, needed for the standardization. Third, the observed and expected numbers are used to calculate the standardized PAR (equation 7) and RR (equation 8).

Estimation of the number of exposed TRALI patients (A in table 1 of the main paper) can be achieved by:

$$A = \sum_{i=1}^{N} A_i = \sum_{i=1}^{N} a_i - n_{1i}p_i + p_i \quad (5.)$$

Which equals the sum of all stratum specific contributions to A (A_i). These contributions are defined by the observed number of transfusions from female donors (i.e. a_i) minus the expected number of female donors in all non-causal transfusions. The expected number of non-causal transfusions from female donors is rewritten as $n_{1i}p_i+p_i$, which equals $(n_{1i}\text{-}1)p_i$ (i.e. the number of non-causal transfusions multiplied by the fraction of female donors). Since his method sums the total number of female donors and subtracts the number of female donors among non-causal transfusions, it effectively sums the number of female donors among causal transfusions (i.e. A in table 1).

Summation of p_i into the weighted average p^* is given by:

$$p^* = \frac{\sum_{i=1}^{N} \frac{(a_i - n_{1i}p_i)}{p_i(1-p_i)} p_i}{\sum_{i=1}^{N} \frac{(a_i - n_{1i}p_i)}{p_i(1-p_i)}} \quad (6.)$$

The expected fraction of exposed cases (p^*) is needed to estimate the observed departure from the expected. The weight should therefore reflect the contribution that each

stratum makes to this departure. The numerator of the weight (a_i-$n_{1i}p_i$) is the actual contribution that each stratum makes to the departure from the expected. The denominator has two properties. First, by dividing by (1-p_i) the weight is corrected for the maximum possible departure this stratum could have contributed, give p_i. Second, dividing by p_i corrects for p_i itself. This gives a weight proportional to the contribution that this stratum would have made to the departure from the expected if everybody in that stratum were exposed. In the event of effect modification this standardization procedure is the most appropriate because the contribution of the different strata among the exposed remains unchanged.[16]

The observed and expected measures defined in equations 5 and 6 can be used to calculate the PAR. This is done by dividing the deviation of the observed from the expected by the maximum possible deviation, as given by:

$$\text{PAR} = \frac{A - N_1 p^*}{N_1 - N_1 p^*} \qquad (7.)$$

Where N_1 is the total number of TRALI cases. By rearranging equation 1 it can be shown that the estimate of the relative risk (RR) defined in terms of the population attributable risk (PAR) can be found as:

$$\text{RR} = \frac{PAR}{(1 - PAR)p^*} + 1 \qquad (8.)$$

To derive confidence intervals for the PAR, the variance of the PAR can be estimated by the variance of A, which is the sum of the variances of a_i (see below). Another option would be to apply bootstrapping procedures. Furthermore, when either p_i or the contributions of each stratum to the PAR are reasonably homogeneous across strata, the normal non-weighted average of p_i can be entered directly into equations 7 and 8, without the need to calculate p^*.

Variance estimation for the PAR

The variance of the standardized PAR, as given in equation 7, can be estimated by:

$$Var(\text{PAR}) = Var\left(\frac{\sum_{i=1}^{N}(a_i - n_{1i}p_i)}{N - \sum_{i=1}^{N} p_i}\right) = \frac{\sum_{i=1}^{N} Var(a_i)}{\left(N - \sum_{i=1}^{N} p_i\right)^2} \qquad (9.)$$

The variance of a_i (i.e. $Var(a_i)$) depends on the probability that the i^{th} case is a type I TRALI, which is given by:

$$\Pr(\text{type I}) = I_i = \frac{(a_i - n_{1i}p_i)}{(1-p_i)} \quad (10.)$$

This probability can be used to calculate the weighted sum of the variance of a_i in the case of a type I TRALI (variance of a binomial(n_{1i}-1, p_i) distribution) and the variance of a_i in the case of a type II TRALI (variance of a binomial(n_{1i}, p_i) distribution). Weighting these variances according to the probability given in equation 10 gives the variance of a_i:

$$Var(a_i) = I_i p_i (n_{1i}-1)(1-p_i) + n_{1i}p_i(1-I_i)(1-p_i) \quad (11.)$$

Correction by maximum likelihood estimation

In this model there are two unknown parameters, which are estimated simultaneously. Firstly, p is the fraction of female donors in the reference population. Secondly, PAR is the fraction of TRALI cases preventable by the exclusion of all female donors.

For a patient with type I TRALI the causal donor is female. The number of female donors among the remaining (n_1-1) donors follows a binomial distribution. The probability that we observe a female donors out of a total of n_1 donors is therefore given by:

$$\Pr(a_i=a|\text{type I}) = \binom{n_1}{a} p^a (1-p)^{n_1-a} \frac{a}{n_1 p} \quad (12.)$$

For a patient with type II TRALI, the total number of female donors will follow a binomial distribution. The probability that we will observe a female donors out of a total of n_1 donors is therefore given by:

$$\Pr(a_i=a|\text{type II}) = \binom{n_1}{a} p^a (1-p)^{n_1-a} \quad (13.)$$

The probability that a given TRALI is of type I is the PAR and the probability that it is of type II is (1-PAR). This yields, for a person for whom the type of TRALI is unknown, that the probability of observing a female donors out of n_1 total donors is:

$$\Pr(a_i=a) = \text{PAR} \times \Pr(a_i=a|\text{type I}) + (1-\text{PAR}) \times \Pr(a_i=a|\text{type II})$$

$$= PAR\binom{n_1}{a} p^a (1-p)^{n_1-a} \frac{a}{n_1 p} + (1-PAR)\binom{n_1}{a} p^a (1-p)^{n_1-a} \quad (14.)$$

The likelihood function for the data of all N TRALI cases is given by:

$$L(p,PAR) = \Pr(a_{i1}=a_1, a_{i2}=a_2, \ldots, a_{iN}=a_N,)$$

$$= \prod_{i=1}^{j} PAR\binom{n_{1i}}{a_i} p^{a_i} (1-p)^{n_{1i}-a_i} \frac{a_i}{n_{1i} p} + (1-PAR)\binom{n_{1i}}{a_i} p^{a_i} (1-p)^{n_{1i}-a_i} \quad (15.)$$

Maximizing this function with respect to PAR and p, yields the maximum likelihood estimates for PAR and p. Numerical methods are needed to maximize this function; we used the function "optim" of the statistical package R.[13] The second derivatives of the log-likelihood function can be used to estimate the standard errors of the estimated parameters, however in this particular case it is easier to calculate profile-likelihood based confidence intervals.[14]