Cover Page

Universiteit Leiden

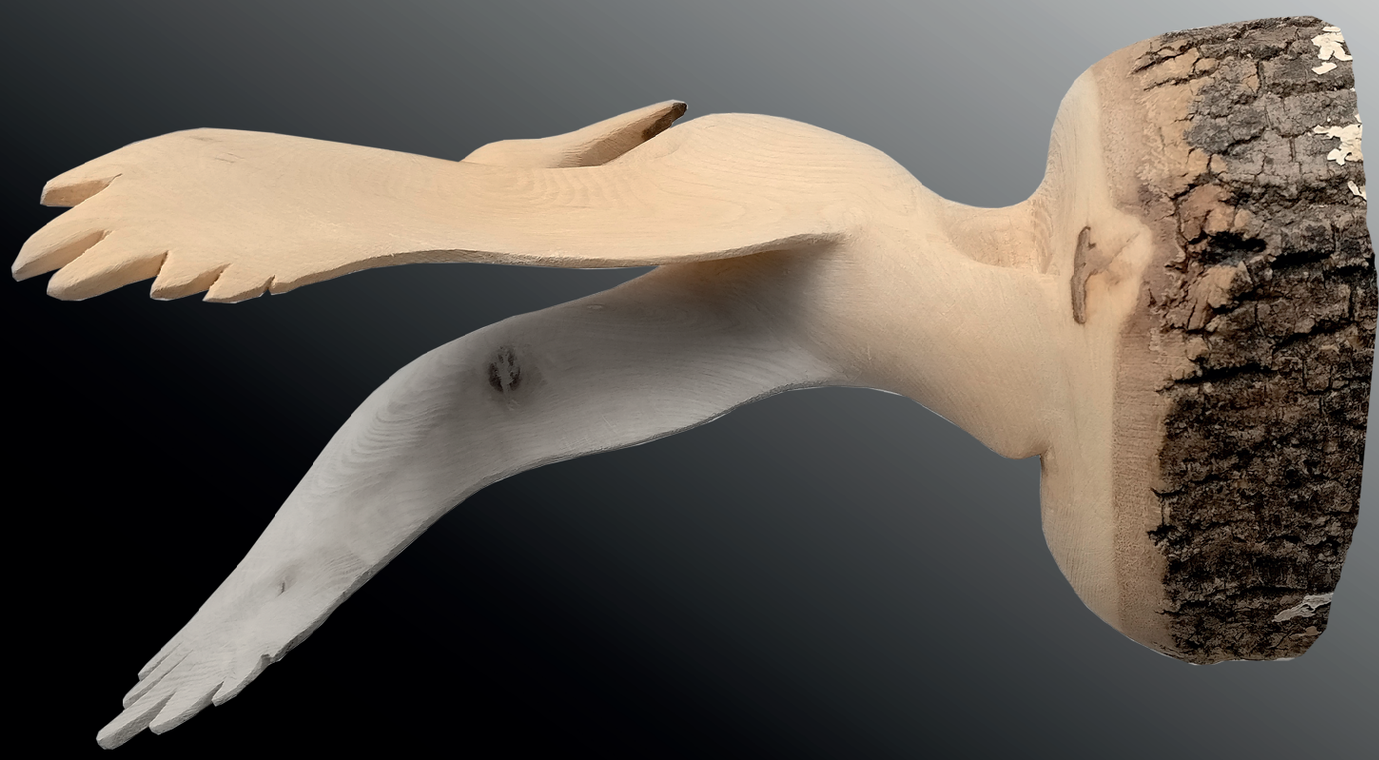
The handle http://hdl.handle.net/1887/55905 holds various files of this Leiden University dissertation.

Author: Caram Deelder, C.
Title: The bright and the dark side of blood transfusion : turning data into knowledge
Issue Date: 2017-11-30

THE BRIGHT AND THE DARK SIDE OF BLOOD TRANSFUSION

TURNING DATA INTO KNOWLEDGE

CAMILA CARAM DEELDER
2017



THE BRIGHT AND THE DARK SIDE OF BLOOD TRANSFUSION

TURNING DATA INTO KNOWLEDGE

Camila Caram Deelder

ISBN: 978-94-6182-837-8
Printing: Off Page, www.offpage.nl
Copyright © 2017 C. Caram Deelder
Cover: Sculpture by Jaap Deelder
Design by Angela Espalaor Titanero and Ana Espalaor Titanero

THE BRIGHT AND THE DARK SIDE OF BLOOD TRANSFUSION

TURNING DATA INTO KNOWLEDGE

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
ter verdediging op 30 november 2017
klokke 16.15 uur

door

Camila Caram Deelder

geboren te Belo Horizonte, Brasil
in 1977

Promotor

Prof. dr. J.G. van der Bom

Co-promotor

Dr. R.A. Middelburg

Leden promotiecommissie

Dr. J. P. Wallis (NHS Foundation Trust, UK)

Dr. S. M. Rezende (Universidade Federal de Minas Gerais, Brazil)

Prof. dr. M. de Haas

Prof. dr. S.C. Cannegieter

The work described in this thesis was performed at the department of Clinical Transfusion Research, Sanquin Research, Leiden and the Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, the Netherlands. The research was fully funded by Sanquin blood supply, division research as part of a full PhD project.

Table of contents

CHAPTER 1

Introduction

PRELUDE

"To his good friends thus wide, I'll ope my arms;
And like the kind life-rendering pelican,
Repast them with my blood."[1]

(The Tragedy of Hamlet. act IV, scene V)

Figure 1 – Sanguine sanguinem sanans. Het Nederlandsche Roode Kruis, 1939.

Centuries before the routine blood transfusion, William Shakespeare refers to blood as a source of life, in the legend of the Dalmatian pelican. This legend tells the history of a mother pelican that with sharp pecks of her beak wounds her own breast, causing rivulets of blood to flow into the mouths of her starving offspring, thus saving their lives. Nowadays the pelican symbolises the altruism of blood donors, who just like the mother pelican give generous donations of blood to help others.

The first attempt to transfuse blood was made in the 17th century shortly after the description of the blood circulation by the English physician William Harvey.[2,3] The first recorded successful human to human transfusion occurred almost 200 years later, in 1818, to treat a woman suffering severe postpartum haemorrhage.[2,4] In 1900 Karl Landsteiner discovered ABO blood groups making it possible to insure blood compatibility and avoid acute ABO transfusion reactions. In that time, blood was transfused as soon as possible after being drawn from donors because no medium was known to postpone blood clotting and allow storage.[4]

After the development of anticoagulants (like sodium citrate), to store and preserve blood, stored blood became a therapeutic possibility. Whole blood could then be stored and kept stocked in glass containers until needed for a transfusion. It also became possible to transport blood over long distances. During the second great war, nationwide programs for blood collection were established in the United States and Britain and blood transfusions were carried out on a large scale for the first time in history.

In the 70's, with the implementation of centrifuge techniques to spin-separate the whole blood into components (i.e. plasma, platelets and leucocytes, and red blood cells – figure 2), and the establishment of plastic bags in replacement to glass bottles, the whole blood therapy gradually gave way to component therapies until, in the 80's, this became the standard of care.[5-8]

Components therapy helped dramatically improve the logistics of transfusion by prolonging storage time, increasing resource use, and decreasing waste. Each component can be stored at its optimal condition and lifespan. Components therapy also allows treatments tailored to specific diseases because it has the advantage of giving patients only what they need. As a consequence, side-effects of components that patients do not need and would have received only because they were present in the 'whole blood' are avoided.[5,8] For example, with component therapy it is possible to transfuse only platelets to hemato-oncological patients, or only red blood cells to sickle cells patients reducing the chances, in both cases, of transfusion-associated circulatory overload (TACO).[6]

Due to the decreased risk of side-effects, components therapy also made prophylactic treatments more feasible. For instance, prophylactic platelet transfusions for children with leukaemia who are treated with high dose chemotherapy, or after stem cell transplantation.[6]

Despite the development of blood component therapy and other technological improvements to produce blood and to ensure its safety, over the last decades, side-effects of transfusions still happen. Short term side effects are routinely monitored and strategies have been implemented to prevent them.[9-14] Conversely, long term side effects of blood transfusions are even now hidden and difficult to uncover. It is important to study these effects and uncover their mechanism to provide clues on how they can be prevented.

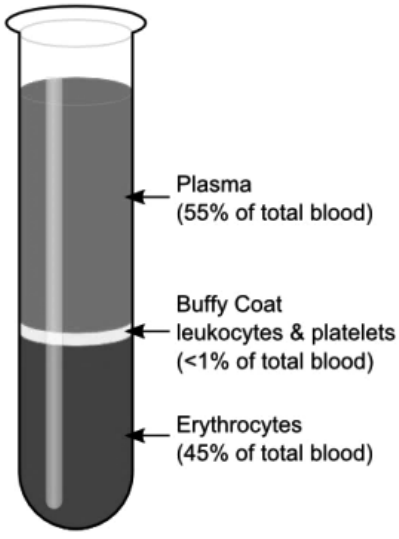


Figure 2 - Components of blood after spin-separation

THE BRIGHT SIDE OF BLOOD TRANSFUSIONS

The usage of blood and the benefits are indisputable. In the last decades blood transfusions had an important role in the treatment of patients. In life threatening situations such as blood haemorrhage and severe anaemia blood is given to stabilise patients and increase recovery speed. Additionally blood transfusions can also be given as prophylactic treatments to prevent further disease complications.[2,6,8,15-17]

Current in the Netherlands the standard blood components are red blood cells, platelets and plasma. Table 1 shows for each component the most common indications, their shelf-life and storage temperature according to Dutch guidelines.[6] Figure 3 shows the total of donations and usage of blood components in the Netherlands, per year, from 2009 to 2015.[18]

Transfusion of red blood cells aims to improve tissue oxygenation because red blood cells are the main transport mechanism for oxygen.[6,8,16] Thus, red blood cells are used to treat patients with a reduced capacity to transport oxygen, such as anaemics (including sickle cell disease) and patients who need acute treatment in case of bleeding.[6,16] Red blood cells can be stored refrigerated for a maximum of 35 days. On average half a million red blood cell units are transfused yearly in the Netherlands.[18]

Platelets as therapeutic intervention are used to treat patients with severe bleeding and thrombocytopenia (deficiency of platelets in the blood). Prophylactic platelets can be transfused prior to a surgical intervention, or for patients at a high perceived bleeding risk including those with combined coagulation deficits, platelet dysfunction due to the use of antiplatelet agents, or renal insufficiency.[6,8,16,19] Platelets are stored in motion at room temperature and can be stored up to 7 days.[6] On average 270,000 platelets buffy coats from whole blood donations are transfused yearly in the Netherlands corresponding to 54,000 units of platelet concentrates pooled from 5 blood donations each.[8,17,18]

Table 1 - Shelf-life, temperature and the indications of blood components.

Component - donation method and production	**Frequent indications**	**Shelf-life and temperature**
Red blood cells (Erythrocytes)	Shortage of oxygen transport	
- whole blood from 1 donor	capacity (bleeding/severe anaemia)	35 days at 2–6 ºC
Platelets		in PAS*-B : 5 days,
- buffy coat from 5 donors		in PAS*-C: and plasma 7 days;
- apheresis from 1 donor	Thrombocytopenia and severe bleeding	at 20–24 ºC in a shaker
Plasma	Deficient clotting factors and	
- apheresis from 1 donor	bleeding combined with clotting	1 donor plasma: 2 years at ≤-25°C;
- apheresis pooled plasma	factor deficiencies	pooled plasma: 4 years at ≤-18°C

PAS: platelet additive solution, B and C refer to the product generation[20]

Plasma is indicated to patients with deficient clotting factors and to treat patients who combine bleeding with deficient clotting factors, resulting from massive blood or volume replacement. It is also indicated for patients with congenital factor deficiencies for which there is no coagulation concentrate available, such as deficiencies of factor V.[6,8] Plasma must to be stored frozen and can be stored as long as 2 years when it is from one donor, and up to four years when plasma is a pooled product from several donors.[6] On average 77,000 units of plasma are transfused yearly in the Netherlands.

Whole blood donations represent 60% of all donations in the Netherlands. In 2015 the total number of donations was over 720,000 from which 285,000 were apheresis donations and 435,000 whole blood donations. Apheresis donations represent the remaining 40% of donations and are mainly plasma.

Apheresis platelets are specifically collected for specific patient profiles and represent approximately 10% of all platelets transfused.[17,18]

In 2009 about 560,000 units of red blood cells, 50,000 units of pooled platelet concentrates and 90,000 units of plasma were transfused to patients. Over the six subsequent years there was a drop of 20% in red blood cell usage and consequently whole blood donations (figure 3). In this period platelets remained stable.[18] This decrease in blood usage follows a trend in which doctors worldwide understand that, despite of the benefits of blood transfusions, more restrictive transfusion strategies than they previously used (i.e. transfusion at a trigger lower than 8 g/dL) are safe in most clinical settings.[21-24]

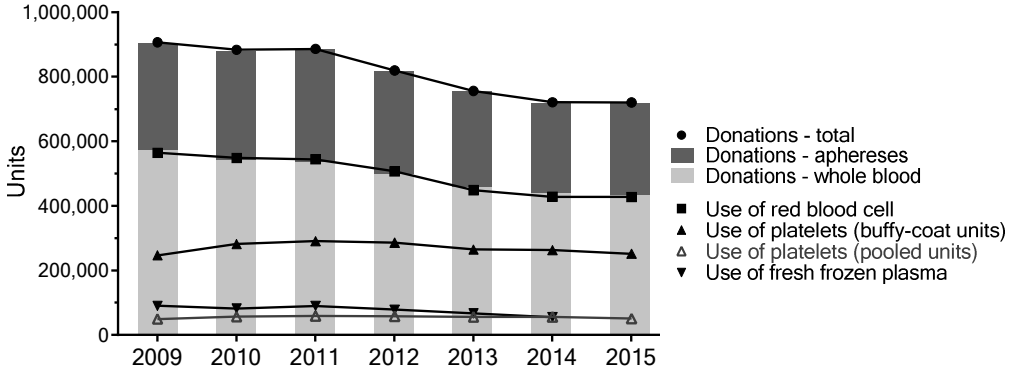


Figure 3 - Blood donations and components used per year in the Netherlands from 2009 to 2015[18]

THE DARK SIDE OF BLOOD TRANSFUSIONS

Blood transfusion as any other medical intervention has side effects. Side effects of blood transfusions are usually called '*transfusion reactions*'. The term transfusion reactions is normally reserved for short terms side effects and includes acute non-infectious complications such as haemolytic transfusion reaction (HTR); febrile non-haemolytic transfusion reaction (FNHTR); transfusion-associated circulatory overload (TACO); and transfusion related acute lung injury (TRALI). These complications cause body changes in the patients and can be seen and measured, for example, shortness of breath, increased temperature or skin rashes. Transfusion reactions also include infectious complications such as post-transfusion viral infections (e.g. HIV, hepatitis and malaria) and transmission of bacterial contamination during production and storage of blood components. These can also be routinely diagnosed through clinical and laboratory tests.[6,9-14] Additionally, blood transfusions can also have other side effects than 'transfusion reactions' which are generally more difficult to be recognised as a side effect of the given transfusion and consequently harder to study.

Three important factors make the study of these other side effects of blood transfusions complex. The first one is that the associations between the transfusion and the outcomes are blurred by events that occur between the time of the transfusion and the outcome. Therefore transfusions are unlikely to be identified as sufficient cause of outcomes. However, studies have shown that transfusions are potential causal components of outcomes such as survival, length of hospital stay, kidney injury, kidney failure or heart failure.[25]

The second factor is *'repeated (or multiple) exposure'*: patients who need blood are likely to need more than one transfusion either in emergency situations during blood loss or in prophylactic treatments over time. Each unit of blood has its own characteristics such as storage time, medium, number of cells in the bag and blood type/compatibility. Each individual transfusion also has its own characteristics such as the clinical status of the patient, baseline disease and disease progression status. All these characteristics have different relationships with outcomes and their effects are mixed over time which makes it difficult (or even impossible) to separate the contribution of each transfusion individually.

The third factor that makes the study of "*long term non-infectious*" side effects of blood transfusions complex is '*confounding by indication*'.[25-27] Confounding by indication is a specific type of confounding that can occur often in non-experimental studies, usually related to the usage of drugs and interventions such as blood transfusions. It occurs because the patients who had blood transfusions have characteristics that made a doctor indicate a certain quantity and blood component. The population of the patient that receives the prescribed amount of blood components is, by definition, different than any other population with a different blood transfusion treatment profile (quantities and blood components, if any). Figure 4 shows two examples of confounding by indication that were previously described in the literature.[28] It shows the relationship between blood transfusion and poor outcome.

In panel A, a spurious association could be created between female donors and poor outcome because patients who receive more transfusions are unlikely to receive exclusively transfusions from male donors. In other words, patients who receive only male transfusions are likely to be less seriously ill and receive less transfusions than patients who were more seriously ill, need more transfusions and consequently receive also transfusions from female donors.

In the panel B a spurious association between young blood and poor outcome is created because clinicians prescribed young components to more seriously ill patients.

If ignored, these 3 factors can lead to wrong conclusions due to bias. When estimates are well adjusted for confounders or the correct sub-population is selected and analysed separately, estimators are more robust.

However, a current problem, especially in observational studies using routinely collected clinical data, is the lack of registration of important variables, resulting in a lack of correction for confounders. In this manner, observational studies, which are the majority of published blood transfusion studies, are often criticized due to their possible biased results.

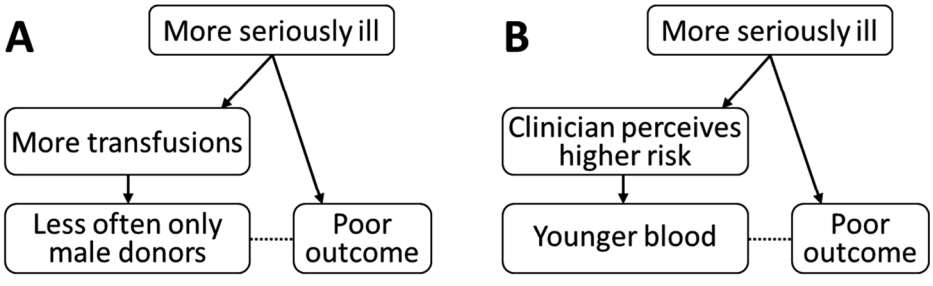


Figure 4 – Cartoons of indirect confounding by indication.[28]

Arrows indicate causal relationships and the dashed lines denote spurious associations.

(A) If transfusions are allocated independent of donor sex, receiving more transfusions will reduce the probability of receiving all transfusions from male donors. Since the number of transfusions was (spuriously) positively associated with poor outcome, receiving all transfusions from male donors will become (spuriously) negatively associated with poor outcome. If in this example we were interested in the relation between donor sex and a negative outcome (e.g., transfusion-related acute lung injury [TRALI]) it would seem like receiving transfusions only from male donors is protecting against TRALI (and therefore like female donors are causing TRALI), while this association is actually not causal.

(B) If clinicians believe younger blood to be safer and therefore specifically reserve or order younger blood for more vulnerable patients with poorer prognosis, a (spurious) negative association between high product age and poor outcome will be created. If in this example we were interested in the relation between storage time and a negative outcome (e.g., mortality) it would seem like receiving younger blood causes mortality, while this association is actually not causal but created by the clinician.

DATA, KNOWLEDGE AND BLOOD TRANSFUSIONS

In science, as in others intellectual activities, data from more than one observation needs to be summarised (i.e. transformed) into information and translated to knowledge to be understood and/or used. The bigger the number of observations and the variability, the more complex these tasks can be. The theoretical model that describes this path is known as the Knowledge Hierarchy or Data-Information-Knowledge Hierarchy. Figure 5 shows a graphical representation of this model in the context of medical science and blood transfusions. This model assumes that data can be used to create information and information can be used to create knowledge[29,30].

Data are facts, or "units of information", i.e. discrete entities without interpretation. Observational clinical studies often use data that were generated for a different purpose, known as secondary data, e.g. patient medical records in a hospital. The data in these sources has a great potential to provide important answers about large patient groups with a wide variety of patients profiles. Observational studies have limitations, mostly related to data quality, which must be addressed in a proper manner.

However, these limitations do not rule out their potential. In other words, also secondary data, generated for other purposes than research can, through scientific methods, be synthetized as information and interpreted as knowledge.[25,29-32]

A variation of the model adds wisdom to the top of the pyramid and refers specific to the *Data-Information-Knowledge-Wisdom Hierarchy* (DIKW). Wisdom is the implementation of knowledge in a form of policies, guidelines and procedures.

These concepts can already be found in transfusion medicine. Its potential is promising and includes patient blood management; benchmarking; patterns of blood use by procedure over time; and detection of transfusion-related complications.[33] A new trend to implement national blood transfusion databases (i.e. to use routinely collected databases structured on national levels) can be seen recently in several countries such as Sweden and Denmark (SCANDAT)[34], the Netherlands[35], Finland[36], England[33] and the multi-national collaboration between the United States of America, Brazil, China, and South Africa.[37]

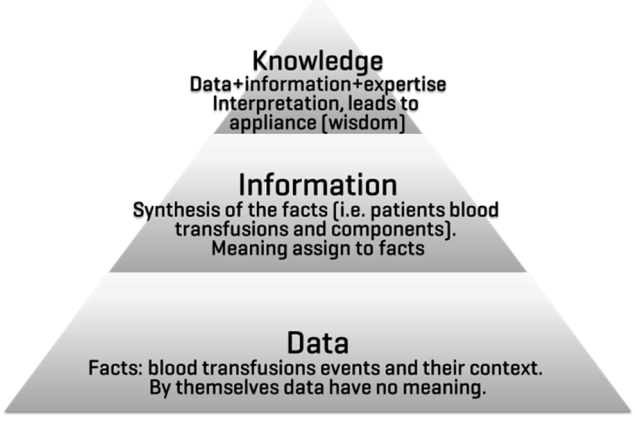


Figure 5 – knowledge hierarchy pyramid and blood transfusion.

These national and multi-national data warehouses are designed to routinely collect and merge information covering the whole blood transfusion chain, from donations to post transfusion events. They can be extremely useful to answer a wide range of transfusion research questions as well as aim to investigate causal relationships. For this purpose they collect, link and anonymise information about:

- Donors
- Donations events
- Products productions process
- Products post production treatments
- Hospitals and transfusion clinics
- Patients
- Pre-transfusion, including indication and pre-transfusion measurements
- Transfusions events
- Post-transfusions, including post transfusions reactions and measurements

Conceptually these warehouses based in daily routine collected data studies (i.e. secondary data studies) and meta-analyses are very closely related.

Both merge complex datasets from various sources and pool relevant data into one estimation.[34,35,38,39] Both are higher in the level of evidence pyramid (figure 5) because they both are closer to no-sampling (i.e. study of populations instead of sampling from populations). Finally, they both rely on advanced algorithms and computer power and became feasible to be performed in the recent decades.

In the same way, using different sources of data, meta-analyses are performed. Meta-analysis is a technique to merge and contrast results (i.e. data) from multiple studies. This consists of identifying patterns or sources of disagreement among the results of studies and when possible summarise the data, as estimations, into information and knowledge.[26] Meta-analyses are the ultimate stage of science and became a strong pillar of the evidence-based medicine. They help doctors update their knowledge without the need of going through the extensive process of searching-judging-quality-reading-summarizing the vast number of papers that science produces nowadays.[40]

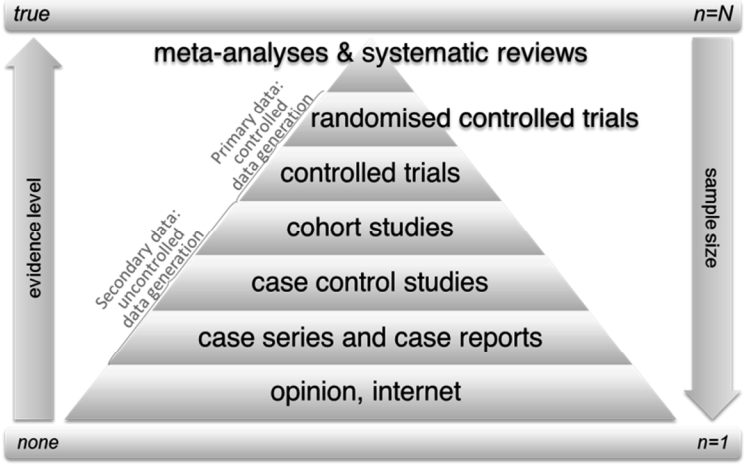


Figure 6 – adapted levels of evidence pyramid
Evidence level increases from bottom to top (left arrow) inversely to sample size (right arrow).
Case series, case report, case control and cohort studies use secondary data (generated in a uncontrolled manner) while non-randomised and randomised controlled trials use primary data (generated in a controlled manner).

OUTLINE

This thesis shows the use of secondary data to produce knowledge in two manners. First as original research (**chapter 2** and **3**), combining daily routine databases of several hospitals. Second as meta-analyses (**chapter 4, 5 and 6)**, combining results of different studies.

In **chapter 2** the relationship between sex of the patient, sex of the donor of red blood cell units and the pregnancy history of female donors was studied. Datasets of 6 different Dutch hospitals during 10 years of total follow up were combined to form a cohort of patients who received blood.

In **chapter 3** an algorithm was used to select a group of patients from a cohort of patients transfused in 10 hospitals based on their specific pattern of platelet transfusions, even when their diagnoses were not available in the source database. Furthermore, for the selected patients the relationship between 'time to the next transfusion' and the 'storage time' of transfused platelet products was investigated.

Chapter 4 and **chapter 5** present two complementary systematic reviews and meta-analyses, addressing the relation between platelet storage time and several outcomes, split in: (i) platelets measurements (**chapter 4**); and (ii) clinical outcomes (**chapter 5**).

Chapter 6 uses the underlying distribution derived from meta-analysed studies to explore the definition of "failed transfusions" in different cut-off points and its relation to components age (fresh versus old platelets).

Finally, a methodological questions was addressed: the use of the terms 'prospective' and 'retrospective' in clinical observational research and its relationship with quality of report (**chapter 7**).

REFERENCES

1. Shakespeare W, Brooke T, Crawford JR. *The tragedy of Hamlet, prince of Denmark*. New Haven; London: Yale University press; G. Cumberlege, Oxford University Press; 1947.
2. AABB. Highlights of Transfusion Medicine History. http://www.aabb.org/tm/Pages/highlights.aspx.
3. Learoyd P. The history of blood transfusion prior to the 20th century--part 1. *Transfusion medicine (Oxford, England)*. 2012;22(5):308-314.
4. Learoyd P. The history of blood transfusion prior to the 20th century--part 2. *Transfusion medicine (Oxford, England)*. 2012;22(6):372-376.
5. Zielinski MD, Jenkins DH, Hughes JD, Badjie KS, Stubbs JR. Back to the future: the renaissance of whole-blood transfusions for massively hemorrhaging patients. *Surgery*. 2014;155(5):883-886.
6. de Vries R, Haas F. English translation of the Dutch Blood Transfusion guideline 2011. *Vox Sanguinis*. 2012;103(4):363.
7. Walter CW. Invention and development of the blood bag. *Vox Sang*. 1984;47(4):318-324.
8. Arya RC, Wander G, Gupta P. Blood component therapy: Which, when and how much. *Journal of anaesthesiology, clinical pharmacology*. 2011;27(2):278-284.
9. Wiersum-Osselton JC. [International databases: usefulness and limitations in hemovigilance]. *Transfusion clinique et biologique : journal de la Societe francaise de transfusion sanguine*. 2010;17(5-6):306-309.
10. Wiersum-Osselton JC, Schipperus MR. [Transfusion reactions in patients: haemovigilance reports to the Dutch National Haemovigilance Office in 2003]. *Nederlands tijdschrift voor geneeskunde*. 2005;149(47):2622-2627.
11. Wiersum-Osselton JC, Faber JC, Politis C, Brand A, van der Bom JG, Schipperus MR. Quality validation of data in national haemovigilance systems in Europe: report of a survey on current state of practice. *Vox Sang*. 2013;104(3):214-217.
12. Politis C, Wiersum JC, Richardson C, et al. The International Haemovigilance Network Database for the Surveillance of Adverse Reactions and Events in Donors and Recipients of Blood Components: technical issues and results. *Vox Sang*. 2016;111(4):409-417.
13. Bolton-Maggs PH. Conference report: the 2015 SHOT symposium and report--what's new? *Transfusion medicine (Oxford, England)*. 2015;25(5):295-298.
14. Bolton-Maggs PH, Cohen H. Serious Hazards of Transfusion (SHOT) haemovigilance and progress is improving transfusion safety. *British journal of haematology*. 2013;163(3):303-314.
15. Fasano RM, Josephson CD. Platelet transfusion goals in oncology patients. *Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program*. 2015;2015:462-470.
16. McIntyre L, Tinmouth AT, Fergusson DA. Blood component transfusion in critically ill patients. *Current opinion in critical care*. 2013;19(4):326-333.
17. van der Meer PF. PAS or plasma for storage of platelets? A concise review. *Transfusion medicine*. 2016;26(5):339-342.
18. Sanquin. Jaarverslag 2015. 2016; http://www.sanquin.nl/over-sanquin/over-sanquin/jaarverslagen/.
19. Rao MP, Boralessa H, Morgan C, et al. Blood component use in critically ill patients. *Anaesthesia*. 2002;57(6):530-534.
20. Ashford P, Gullikkson H, Georgsen J, Distler P. Standard terminology for platelet additive solutions. *Vox Sang*. 2010;98(4):577-578.
21. Holst LB, Petersen MW, Haase N, Perner A, Wetterslev J. Restrictive versus liberal transfusion strategy for red blood cell transfusion: systematic review of randomised trials with meta-analysis and trial sequential analysis. *BMJ (Clinical research ed)*. 2015;350:h1354.
22. Kranenburg FJ, Arbous MS, So-Osman C, van der Bom JG. [Blood is not for everyone: the usefulness of erythrocyte transfusion]. *Nederlands tijdschrift voor geneeskunde*. 2015;159:A9408.
23. Whitaker B, Rajbhandary S, Kleinman S, Harris A, Kamani N. Trends in United States blood collection and transfusion: results from the 2013 AABB Blood Collection, Utilization, and Patient Blood Management Survey. *Transfusion*. 2016;56(9):2173-2183.
24. Carson JL, Terrin ML, Noveck H, et al. Liberal or restrictive transfusion in high-risk patients after hip surgery. *The New England journal of medicine*. 2011;365(26):2453-2462.
25. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
26. Rothman KJ. *Epidemiology: An Introduction*. OUP USA; 2012.
27. Stafford KA, Boutin M, Evans SR, Harris AD. Difficulties in demonstrating superiority of an antibiotic for multidrug-resistant bacteria in nonrandomized studies. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2014;59(8):1142-1147.
28. Middelburg RA, van de Watering LM, van der Bom JG. Blood transfusions: good or bad? Confounding by indication, an underestimated problem in clinical transfusion research. *Transfusion*. 2010;50(6):1181-1183.
29. Chaffey D, Wood S. *Business Information Management: Improving Performance Using Information Systems*. Financial Times Prentice Hall; 2005.
30. Rowley J. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*. 2007;33(2):163-180.
31. Mues KE, Liede A, Liu J, et al. Use of the Medicare database in epidemiologic and health services research: a valuable source of real-world evidence on the older and disabled populations in the US. *Clinical Epidemiology* 2017 (9):267–277.
32. Ronquillo C, Currie LM, Rodney P. The Evolution of Data-Information-Knowledge-Wisdom in Nursing Informatics. *ANS Advances in nursing science*. 2016;39(1):E1-18.
33. Pendry K. The use of big data in transfusion medicine. *Transfusion medicine (Oxford, England)*. 2015;25(3):129-137.
34. Edgren G, Rostgaard K, Vasan SK, et al. The new Scandinavian Donations and Transfusions database (SCANDAT2): a blood safety resource with added versatility. *Transfusion*. 2015;55(7):1600-1606.
35. van Hoeven LR, Hooftman BH, Janssen MP, et al. Protocol for a national blood transfusion data warehouse from donor to recipient. *BMJ open*. 2016;6(8):e010962.
36. Palo R, Ali-Melkkilä T, Hanhela R, et al. Development of permanent national register of blood component use utilizing electronic hospital information systems. *Vox Sanguinis*. 2006;91(2):140-147.
37. Kleinman S, Busch MP, Murphy EL, Shan H, Ness P, Glynn SA. The National Heart, Lung, and Blood Institute Recipient Epidemiology and Donor Evaluation Study (REDS-III): a research program striving to improve blood donor and transfusion recipient outcomes. *Transfusion*. 2014;54(3 Pt 2):942-955.
38. Johnson BT. Meta-Analysis: The Original 'Big Data'. 2013; http://meta-analysis.ning.com/profiles/blogs/meta-analysis-the-original-big-data. Accessed 06/01/2017.
39. Sim I. Two Ways of Knowing: Big Data and Evidence-Based Medicine. *Annals of internal medicine*. 2016;164(8):562-563.
40. Sauerland S, Seiler CM. Role of systematic reviews and meta-analysis in evidence-based medicine. *World journal of surgery*. 2005;29(5):582-587.

CHAPTER 2

Association of blood transfusion from female donors with and without a history of pregnancy with mortality among male and female transfusion recipients

JAMA. 2017; 318(15): 1471-1478.

Camila Caram-Deelder[1;2]; Aukje L. Kreuger[1;2]; Dorothea Evers[1;3]; Karen M.K. de Vooght[4]; Daan van de Kerkhof[5]; Otto Visser[6]; Nathalie C.V. Péquériaux[7]; Francisca Hudig[8]; Jaap Jan Zwaginga[1;3]; Johanna G. van der Bom[1;2]; Rutger A. Middelburg[1,2].

1. Center for Clinical Transfusion Research, Sanquin Research; 2. Department of Clinical Epidemiology, Leiden University Medical Center; 3. Department of Immunohaematology and Blood Transfusion, Leiden University Medical Centre; 4. Department of Clinical Chemistry and Haematology, University Medical Center Utrecht; 5. Department of Clinical Chemistry and Haematology, Catharina Hospital; 6. Department of Haematology, VU Medical Center; 7. Department of Clinical Chemistry and Haematology, Jeroen Bosch Hospital; 8. LabWest, Haga Teaching Hospital.

ABSTRACT

Importance: Transfusion of red blood cells from female donors has been associated with increased mortality of male patients.

Objective: To quantify the association between red blood cell transfusion from female donors with and without a history of pregnancy and mortality of red blood cell recipients.

Design, Setting, and Participants: Retrospective cohort study of first time transfusion recipients at six major Dutch hospitals enrolled from 30/05/2005 to 01/09/2015; the final follow-up date was 01/09/2015. The primary analysis was the no-mixture cohort (i.e. either all transfusions from male donors, or all from female donors without a history of pregnancy, or all from female donors with a history of pregnancy). The association between mortality and exposure to transfusions from ever-pregnant or never-pregnant female donors was analyzed using life tables and time-varying Cox proportional hazards models.

Exposure: Red blood cell transfusions from ever-pregnant or never-pregnant female donors, compared to red blood cell transfusions from male donors.

Main outcomes and measures: All-cause mortality during follow-up.

Results: The cohort for the primary analyses consisted of 31,118 patients (median 65 (IQR 42 to 77) years old; 52% female) who received 59,320 red blood cell transfusions exclusively from one of three types of donors (88% men; 6% ever-pregnant women; and 6 % never-pregnant women). The number of deaths in this cohort was 3,969 (13% mortality). For male red blood cell transfusion recipients, all-cause mortality rates after a red blood cell transfusion from an ever-pregnant female donor versus male donor were 101 versus 80 deaths per 1,000 person years (py), time-dependent "per transfusion" hazard ratio (HR) for death was 1.128 (95% confidence interval (CI): 1.009 to 1.260). For receipt of transfusion from never-pregnant female donor versus male donor, mortality rates were 78 versus 80 deaths per 1,000py, HR 0.928 (CI: 0.809 to 1.064;). Among female red blood cell transfusion recipients, mortality rates for an ever-pregnant female donor versus male donor were 74 versus 62/1,000py, HR 0.993 (CI: 0.870 to 1.133) and for a never-pregnant female donor versus male donor, mortality rates were 74 versus 62/1,000py, HR 1.007 (CI: 0.882 to 1.149).

Conclusions and relevance: Among patients who received red blood cell transfusions, receipt of a transfusion from an ever-pregnant female donor compared to a male donor was associated with increased all-cause mortality among male patients but not among female patients. Transfusions from never-pregnant female donors were not associated with increased mortality among male or female recipients. Further research is needed to replicate these findings, determine their clinical significance, and identify the underlying mechanism.

KEY POINTS

Question: Is there an association between red blood cell transfusion from female donors with and without a history of pregnancy and recipient mortality?

Findings: In this retrospective cohort study that included 31,118 patients who received red blood cell transfusions, receipt of a transfusion from an ever-pregnant female donor was associated with a statistically significant increase in all-cause mortality among male red blood cell transfusion recipients (hazard ratio 1.128) but not among female recipients (hazard ratio 0.993).

Meaning: Receipt of red blood cell transfusion from female donors with a history of pregnancy was associated with increased mortality among male patients. Further research is needed to replicate these findings, determine their significance, and define the underlying mechanism.

INTRODUCTION

Transfusion of red blood cells is among the most commonly performed procedures in hospitals.[1] It has been reported that mortality was increased after transfusion of red blood cells from female donors compared to male donors.[2-7] The most common cause of transfusion related mortality is transfusion related acute lung injury (TRALI), which has also been shown to be associated with transfusions from female donors.[8-10] Furthermore, TRALI is associated specifically with transfusions from female donors with a history of pregnancy.[11,12] This raises the question whether the increased mortality after red blood cell transfusions could also depend on a history of pregnancy of the donor. However, for TRALI it has been shown that only plasma rich products confer a pregnancy related antibody mediated risk, whereas red blood cells do not.[10,11] The increased mortality in recipients of red blood cells from female donors may be related to either immunological phenomena or other mechanisms.

Any proposed immunological mechanism is likely to be dependent on a history of pregnancy of the donor. An absence of association with pregnancy status of the donor would suggest other, non-immunological mechanisms to be more likely. Therefore, the aim of the current study was therefore to quantify the association between red blood cell transfusion from female blood donors, with and without a history of pregnancy, and patient mortality in female and male transfusion recipients.

METHODS

Study design

As previously described, a retrospective cohort of first ever transfusion recipients, transfused from 30/05/2005 to 01/09/2015 in six major Dutch hospitals, was established.[13,14] All patients included in a previous study of mortality after transfusion of red blood cells from female donors were excluded from the current analyses, to create an independent cohort.[2] Ethical approval for this study was obtained from the institutional review board of the Leiden University Medical Center, and local review boards of all participating centers. The review boards waived the need for informed consent, because only routinely collected data were processed after coding to remove identifying information.

Population

Primary analyses were performed in a "no-mixture" cohort, to avoid dilution of effects by mixing comparing patients who received transfusions from both male and female donors. This cohort consisted of patients who received all their red blood cells exclusively from male donors or who received all their red blood cells exclusively from female donors without a history of pregnancy (never-pregnant donors), or who received all their red blood cells exclusively from female donors with a history of pregnancy (ever-pregnant donors). Follow-up time was censored at the time they violated these inclusion criteria. This censoring could occur at time 0, in which case patients contributed 0 follow-up time and were not included in the denominator. Similarly, a "single transfusion" cohort also was selected, consisting of patients who received only a single transfusion. Additionally, all analyses were repeated in the full cohort, to check whether any observed association potentially depended on the selection of the no-mixture cohort. The ethnicity of patients and donors was not recorded.

Recorded data

Patient data

Dates of birth, dates of death and sex of patients, transfusion dates, product types and identification codes of transfused red blood cells were provided by the hospitals from electronic records of the blood transfusion services. All transfusions, given for any indication, were included. Mortality data were verified by the hospitals until the date of data extraction. Mortality data were considered to be complete due to the use of a nationally linked computer system and the legal requirement for reporting all deaths to this system. Therefore, follow-up is considered to be complete; the final follow-up date was 01/09/2015.

Donor and blood product data

Dates of birth, sex, and pregnancy before donation (see supplement for details) were provided by Sanquin (the national Dutch blood supply) and linked to patients' data using the product identification codes of transfused red blood cells. All blood products in the Netherlands are leukocyte-depleted by pre-storage filtration and nearly all products are transfused ABO-Rhesus D identically.

Pregnancy of female blood donors

At their first donation, female blood donors self-reported any previous pregnancy. At all subsequent donations, they reported whether they had been pregnant since the previous donation. However, since some female donors had their first ever donation prior to the establishment of the current electronic recording system at the Sanquin blood bank, the answer to the question at first donation could be missing. When the first donation was registered and answered as never-pregnant the pregnancy status was considered never-pregnant until the first subsequent donation at which a pregnancy was reported. If the first donation was missing, the pregnancy status was considered unknown until the first subsequent donation at which a pregnancy was reported.

Missing data

Information about donors' pregnancy history was not specifically recorded and was therefore missing for 44% of donations from female donors (table 2s). However, missingness depended solely on logistic factors (i.e. changes in the electronic recording of donor information over the years). These data were therefore expected to be "missing completely at random" (as also shown in table 3s), allowing a valid "complete case" analysis.[15] We therefore selected only cases with complete data available.

Statistical analyses

All statistical analyses were performed in Stata 14.1 and pre-specified in the protocol, unless otherwise indicated.[16] The only outcome assessed was all-cause mortality, at any time during follow-up, as specified per participating center in the supplement (table 1s).

Survival analyses were performed with follow-up starting on the day of the first red blood cell transfusion. Follow-up ended at death, or on the reference day, determined for each hospital separately (see supplement table 1s). The reference day was the last day for which the hospital had provided data. Follow-up time of patients in the different cohorts was censored at the time they first violated the inclusion criteria for that cohort. To increase homogeneity, follow-up time of patients who received more than 15 transfusions was censored at the time of the 16th transfusion. All analyses were stratified by patients' sex. Transfusions of other blood products were ignored, because they were not correlated with sex and pregnancy history of the donor of red blood cells (table 3s). All reported p-values are 2-sided and p-values <0.05 were considered statistically significant. No adjustments for multiple comparisons were performed.

Kaplan-Meier

Kaplan-Meier curves were constructed for the "single transfusion" cohort. The analyses were limited to three years of follow-up. At this time differences in cumulative incidence between different groups and 95% confidence intervals for these differences were calculated according to standard formula's (see supplement for more details).

Time varying Cox proportional hazard models

Cox proportional hazard models, including both time-varying and fixed variables, were fitted to correct for potential confounding. All confounding variables (i.e. center (fixed), patients' ABO-Rhesus D blood group (fixed), age of the donor (time-varying), cumulative number of transfusions (time-varying), calendar year (time-varying), and an interaction term for center and number of transfusions (time-varying)) were included in the models as categorical variables, with as many categories as there were exposure levels (see supplement for details on potential confounders).

For the time-varying analyses values of variables could change on each day with red blood cell transfusion(s). At each day with red blood cell transfusion(s) the cumulative number of red blood cell transfusions and of red blood cell transfusions from male, female never-pregnant, and female ever-pregnant donors, up to and including that day were determined.

Exposures (i.e. cumulative number of transfusions from (never/ever-pregnant) female donors (time varying)) were included in the models as continuous variables. Consequently, hazard ratios should be interpreted on a multiplicative scale. However, since the model estimates the hazard ratio based on observed numbers of transfusions only, the hazard ratios should not be extrapolated beyond the observed mean number of transfusions in each cohort (see supplement: table 3s, for an illustration of this interpretation). The proportional hazards assumption was checked for all models and no gross violations of this assumption were detected, implying the hazard ratio can be interpreted as a valid estimate of the average hazard ratio over the observed period.

Separate models were run for the two different exposures (i.e. never-pregnant and ever-pregnant). For the no-mixture cohort this meant exclusion of patients who received any transfusions from the other exposure group, any transfusions with unknown pregnancy history, or a mixture of exposed (i.e. ever-pregnant or never-pregnant, depending on the analyses) and unexposed (i.e. male) units. This way the exposure group of interest was always compared directly to male donors, since all other units were excluded. For the full model recipients of transfusions both from the exposure group of interest and from male donors were additionally included.

Effect measure modification

We previously reported effect measure modification by age of the transfused patients.[2] A primary objective of this study therefore was to also repeat these analyses after stratification by age of the patient for pre-specified categories of age (0-17, 18-50, 51-70, and ≥71 years). Effect measure modification was formally quantified by adding interaction terms for age (p-value for interaction-trend across four categories, from a Z-distribution using standard errors estimated from the observed information matrix) and sex of the patient to the final model.

RESULTS

Population

A total of 42,132 patients received 106,641 units red blood cells, 76% from male donors, 12% from ever-pregnant donors, 12% from never-pregnant female donors. The median number of transfused units per patient was 2 (interquartile range (IQR): 2 to 3). These patients were followed for a median of 380 days (IQR: 27 to 1217), had a median age of 66 years (IQR: 46 to 77), and 21,915 (52%) of them were female. The number of deaths was 6,975 (17%). Among this "full cohort" 31,118 patients received 59,320 units of red blood cells exclusively from one of the three types of donor (i.e. the "no-mixture" cohort: either all units exclusively from male donors, or all units exclusively from female donors without a history of pregnancy (never-pregnant donors), or exclusively from female donors with a history of pregnancy (ever-pregnant donors)). These patients were followed for a median of 245 days (IQR: 9 to 1,172), had a median age of 65 years (IQR: 42 to 77), and 16,123 (52%) of them were female. The number of deaths in the "no-mixture cohort was 3,969 (13%). Table 1 shows a comparison of patient characteristics between the "full cohort", "no-mixture" cohort, and "single transfusion" cohort stratified by patient sex.

Donors' pregnancy history and mortality after red blood cell transfusions

Primary analyses: the "no-mixture" cohort

The hazard ratio for death after one additional unit of red blood cells from a never-pregnant female donor, compared to a unit from a male donor, was 0.928 (CI: 0.809 to 1.064) for male patients and 1.007 (CI: 0.882 to 1.149) for female patients (table 2). The hazard ratio for death after one additional unit of red blood cells from an ever-pregnant female donor, compared to a unit from a male donor, was 1.128 (CI: 1.009 to 1.260) for male patients and 0.993 (CI: 0.870 to 1.133) for female patients (table 2).The highest hazard ratios for death after transfusion of red blood cells from ever-pregnant female donors was observed in male patients under 50 years of age (table 3).

Secondary analyses I: the "single transfusion" cohort

The 3 years cumulative incidence of death among male patients was 13.5% after a transfusion from a male donor, 13.1% after a transfusion from a never-pregnant female donor (difference: 0.4% (CI: -3.8 to 3.0%)) and 16.9% after a transfusion from an ever-pregnant female donor (difference: 3.5% (CI: -0.3 to 7.2%)) (figure 1).

The cumulative incidence of death among female patients was 12.6% after a transfusion from a male donor, 12.0% after a transfusion from a never-pregnant female donor (difference: 0.6% (CI: -3.7 to 2.6%)) and 15.9% after a transfusion from an ever-pregnant female donor (difference: 3.3% (CI: -0.5 to 7.1%)) (figure 1).

Secondary analyses II: the full cohort

The hazard ratio for death after one additional unit of red blood cells from an ever-pregnant female donor, compared to a male donor, was 1.082 (CI: 1.015 to 1.152) for all male patients, 1.178 (CI: 0.824 to 1.685) for male patients aged 0 to 18 years, and 1.432 (CI: 1.126 to 1.823) for male patients aged 18 to 50 years (table 3). For female patients the hazard ratio for death after one additional unit of red blood cells from an ever-pregnant female donor, compared to a male donor, was 0.994 (CI: 0.928 to 1.065).

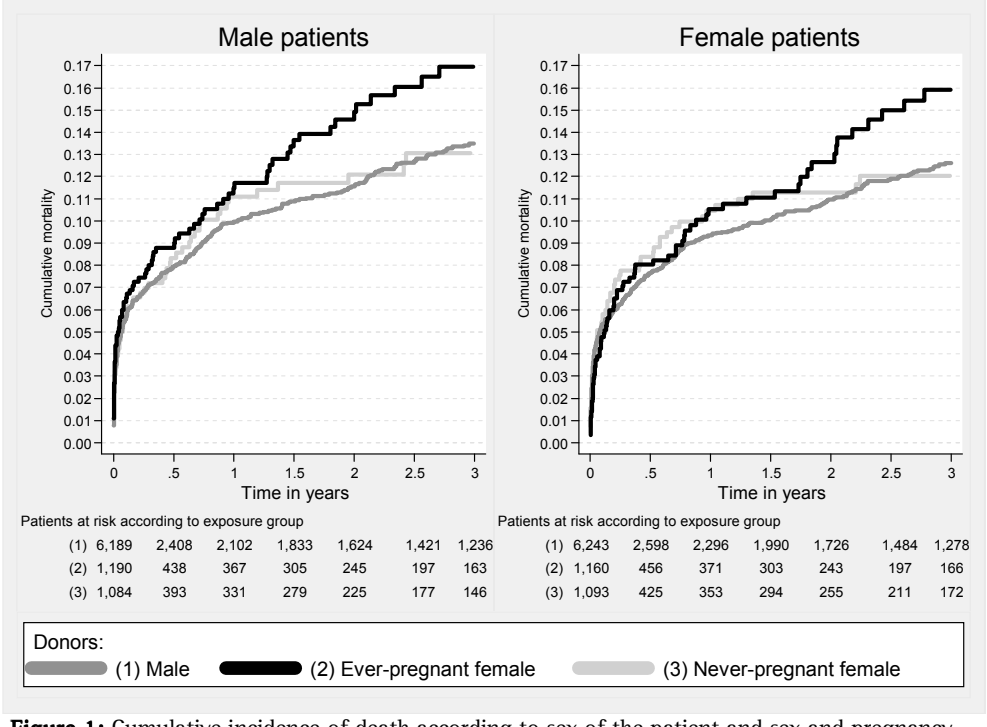


Figure 1: Cumulative incidence of death according to sex of the patient and sex and pregnancy history of the donor in the single transfusion cohort

The number of patients remaining at risk of death in each exposure category is indicated below the x-axis. Single transfusion cohort: consists of all the follow-up time during which patients had received only a single transfusion.
Follow-up time was censored at the time they violated this inclusion criteria.

Table 1: Characteristics of female and male patients and red cells transfusions in the different cohorts

		No-mixture cohort		Single transfusion cohort		Full cohort	
		Male Patients	Female Patients	Male Patients	Female Patients	Male Patients	Female Patients
Number of patients		14,995	16,123	8,463	8,496	20,217	21,915
Number of deaths		1,982 (13%)	1,987 (12%)	606 (7%)	590 (7%)	3,597 (18%)	3,378 (15%)
Follow-up (in days)*		144 (6 to 1,041)	351 (14 to 1,275)	15 (1 to 717)	22 (1 to 781)	274 (18 to 1,106)	479 (44 to 1,316)
Person time (sum in years)		24,339	31,637	10,266	10,971	35,281	45,904
Age of patients (in years)*		65 (45 to 75)	66 (39 to 78)	63 (3 to 75)	64 (21 to 78)	66 (50 to 75)	66 (43 to 78)
0 to 17		2,813 (19%)	2,303 (14%)	2,534 (30%)	2,085 (25%)	2,986 (15%)	2,439 (11%)
18 to 50		1,427 (10%)	2,873 (18%)	603 (7%)	1,073 (13%)	2,140 (11%)	4,246 (19%)
51 to 70		5,244 (35%)	4,253 (26%)	2,406 (28%)	1,966 (23%)	7,460 (37%)	6,112 (28%)
≥71		5,511 (37%)	6,694 (42%)	2,920 (35%)	3,372 (40%)	7,631 (38%)	9,118 (42%)
ABO Rhesus D	**O+**	5,426 (36%)	5812 (36%)	2,945 (35%)	2,854 (34%)	7,366 (36%)	7,874 (36%)
	O-	932 (6%)	1024 (6%)	502 (6%)	562 (7%)	1,279 (6%)	1,388 (6%)
	A+	4,977 (33%)	5,416 (34%)	2,670 (32%)	2,699 (32%)	6,805 (34%)	7,479 (34%)
	A-	930 (6%)	1,051 (7%)	476 (6%)	548 (6%)	1,286 (6%)	1,433 (7%)
	B+	1,240 (8%)	1,361 (8%)	700 (8%)	751 (9%)	1,698 (8%)	1,952 (9%)
	B-	185 (1%)	220 (1%)	99 (1%)	110 (1%)	250 (1%)	318 (1%)
	AB+	433 (3%)	452 (3%)	249 (3%)	251 (3%)	611 (3%)	647 (3%)
	AB-	75 (1%)	86 (1%)	50 (1%)	45 (1%)	114 (1%)	115 (1%)
Not defined		797 (5%)	701 (4%)	772 (9%)	676 (8%)	808 (4%)	709 (3%)
Transfusions of red blood cells units per patient*		2 (1 to 2)	2 (1 to 2)	1 (1 to 1)	1 (1 to 1)	2 (1 to 3)	2 (2 to 3)
Total number of red blood cells units transfused		28,663	30,657	8,463	8,496	52,637	54,004
from male donors		25,190 (88%)	27,155 (89%)	6,189 (73%)	6,243 (73%)	40,022 (76%)	41,267 (76%)
from female ever-pregnant donors		1,818 (6%)	1,810 (6%)	1,190 (14%)	1,160 (14%)	6,528 (12%)	6,454 (12%)
from female never-pregnant donors		1,655 (6%)	1,692 (6%)	1,084 (13%)	1,093 (13%)	6,087 (12%)	6,283 (12%)

Numbers represent numbers of patients (%) or transfusions (%), unless otherwise indicated. * median and interquartile range. No-mixture cohort: consists of all the follow-up time during which patients either received all their red blood cells exclusively from male donors or while they received all their red blood cells exclusively from female donors without a history of pregnancy (never-pregnant donors), or exclusively from female donors with a history of pregnancy (ever-pregnant donors). Single transfusion cohort: consists of patients with only a single red blood cell transfusion during the time period they were followed. Follow-up time was censored at the time they violated these inclusion criteria.

Additional results (supplement)

Cumulative incidences of death, in the single transfusion cohort, at different follow-up times, are shown in figure 3s and table 5s. Table 3s shows the distribution of donor types according to patient sex and plasma and platelets transfusions received. Data on numbers of patients, transfusions, deaths per subgroup – also for all female donors combined, regardless of pregnancy history – are shown in tables 2s, 6s-8s. Results of analyses of red blood cells corrected for plasma and platelet transfusions are shown in tables 9s-10s. Results of analyses for female donors with unknown pregnancy history are shown in table 11s and figure 4s. A direct comparison between ever-pregnant and never-pregnant female donors is shown in table 12s. Analyses of platelet transfusions are shown in the tables 13s-14s.

Effect measure modification

The tests for interaction for the association between transfusion of red blood cells from ever-pregnant donors vs male donors and mortality among male vs female recipients regardless of recipient age did not meet statistical significance (p=0.304 for interaction for the no-mixture cohort, p=0.536 for the single-transfusion cohort and p=0.578 for the full cohort). The strength of the association of ever-pregnant donors and mortality of male patients was different for patients of different ages, as indicated by the p-value for interaction-trend (table 3).

Similarly the differences between male and female patients in the strength of association of ever-pregnant donors with mortality of patients under 50 years of age were statistically significant (p=0.03 for interaction for the no-mixture cohort, p=0.01 for the single-transfusion cohort and p=0.01 for the full cohort).

DISCUSSION

The results from this large retrospective cohort study suggest that the association of female donors with increased mortality among male patients was related to the pregnancy history of female blood donors and the age of the patient. Men who received a red blood cell transfusion from an ever-pregnant female donor had a statistically significant increase in mortality compared to men who received a red blood cell transfusion from a male donor or from a female donor without a history of pregnancy. There was no significant association between pregnancy status of female red blood cell transfusion donors and mortality among female recipients of red blood cell transfusions.

The association of increased mortality among men who received transfusions from ever-pregnant donors suggests a possible mechanism based on immunological changes occurring during pregnancy. Of all changes occurring during pregnancy, the immunological ones are the most enduring. An alternative explanation could be a difference in iron status between (ever-pregnant) female and male donors. Iron deficiency in donors has recently been shown to be associated with worse recovery of red blood cells in recipients in a murine model.[17] Some studies also report differences in red blood cell physiology between the sexes.[13-19]

Results from studies on the association of donor sex and patient mortality, including the current one, tend to be consistent in showing associations for male patients, but not for female patients.[2-6] This specificity for male recipients seems difficult to explain based on differences in red blood cell physiology, supporting a possible role for a sex-specific immunological mechanism. It is difficult to predict whether the small amount of plasma in red blood cell transfusions contains enough antibodies to confer an increased risk of mortality, but it cannot be ruled out.

Furthermore, leukocyte-depleted red blood cell transfusions routinely contain less than a million leukocytes. However, to allow for naturally occurring variation, quality control standards allow up to five million leukocytes in a small percentage of products. These could include both antigen specific lymphocytes or regulatory T-cells.

Some differences exist between results from reported studies on the association between donor sex and recipient mortality.[2-6] These differences could, in the light of the current results, potentially be explained by a combination of differences in prevalence of a history of pregnancy among donors and differences in age distribution of recipients.

This study has several strengths. The large size of the cohort allowed selection of the "no-mixture" cohort, and enabled study of patients who received blood transfusions from only one type of donor (i.e. male vs. previously pregnant female vs. never pregnant female). However, the selection of a no-mixture cohort, could limit generalizability . The patients in the no-mixture cohort receive fewer transfusions, since the probability of receiving mixed transfusions increases with the total number of transfusions. Similarly, the censoring of patients who received 16 or more transfusions could limit generalizability to this group.

This study also has several limitations. First, the difference in effect size and direction between male and female recipients was not significant among recipients of all ages, only among those 50 years and younger. This makes the findings very tentative, and they require validation in other studies Second, this study was retrospective, and data were recorded for routine clinical practice and not specifically for this study. This could cause both inaccuracy of data and unavailability of data. Third, there were missing data particularly regarding pregnancy status for the woman donating red blood cells. Forth, information on cause of death was not available. Fifth, there may have been residual confounding or confounding by an unidentified variable. Sixth, the analysis included a large number of comparisons, but there was no adjustment for multiple comparisons.

Conclusion

Among patients who received red blood cell transfusions, receipt of a transfusion from an ever-pregnant female donor compared to a male donor was associated with increased all-cause mortality among male patients but not among female patients. Transfusions from never-pregnant female donors were not associated with increased mortality among male or female recipients. Further research is needed to replicate these findings, determine their clinical significance,andidentifytheunderlyingmechanism.

SUPPLEMENTAL MATERIAL

Available at: https://goo.gl/TWSkpY

- Methodological details
- Additional Results
- Models corrected for plasma and platelets products received
- Female donors with unknown pregnancy status
- Female never pregnant donors versus female ever pregnant donors
- Other blood components

ACKNOWLEDGEMENTS

We thank Bert Mesman and Herman Geerlings (Sanquin Amsterdam) for data on the Dutch donor population.

Table 2: Mortality hazard ratio of male and female patients exposed to red blood cell transfusions from female (never-pregnant or ever-pregnant) donors in comparison to red blood cell transfusions from male donors in the single transfusion, no-mixture, and full cohorts

Donor Category	Male Recipients			Female Recipients		
	No. of Deaths Among Recipients/Total No.	HR (95% CI)	*p*-value	No. of Deaths Among Recipients/Total No.	HR (95% CI)	*p*-value
No-Mixture Cohort						
Male	1722/12 212	1 [Reference]		1752/13 332	1 [Reference]	
Ever-pregnant female	1873/13 669	1.128 (1.009 to 1.260)	0.034	1871/14 770	0.993 (0.870 to 1.133)	0.916
Never-pregnant female	1722/12 212	0.928 (0.809 to 1.064)	0.285	1868/14 685	1.007 (0.882 to 1.149)	0.923
Single-Transfusion Cohort						
Male	434/6189	1 [Reference]		433/6243	1 [Reference]	
Ever-pregnant female	532/7379	1.225 (0.977 to 1.537)	0.079	517/7403	1.115 (0.878 to 1.417)	0.370
Never-pregnant female	508/7273	0.959 (0.738 to 1.246)	0.754	506/7336	1.002 (0.771 to 1.302)	0.989
Full Cohort						
Male	2538/15 304	1 [Reference]		2448/16 617	1 [Reference]	0.868
Ever-pregnant female	2689/16 338	1.082 (1.015 to 1.152)	0.015	2567/17 654	0.994 (0.928 to 1.065)	
Male	2521/15 163	1 [Reference]		2447/16 608	1 [Reference]	0.234
Never-pregnant female	2630/16 091	1.063 (0.993 to 1.136)	0.077	2563/17 593	0.957 (0.889 to 1.029)	

Numbers represent hazard ratios per transfused unit (95% confidence interval) compared to receiving a unit from a male blood donor. *Reference: male blood donors. All models are adjusted for calendar year, blood group (ABO-RhD), hospital, age of the donor, cumulative number of transfusions and an interaction term for hospital and cumulative number of transfusions. No-mixture cohort: consists of all the follow-up time during which patients either received all their red blood cells exclusively from male donors or while they received all their red blood cells exclusively from female donors without a history of pregnancy (never-pregnant donors), or exclusively from female donors with a history of pregnancy (ever-pregnant donors). Single transfusion cohort: consists of participants who received only a single red blood cell transfusion during the time period of follow-up. Follow-up time was censored at the time they violated these inclusion criteria.

Table 3: Mortality hazard ratio of male patients exposed to red blood cell transfusions from female ever-pregnant donors in comparison to red blood cell transfusions from male donors in the single transfusion, no-mixture, and full cohorts, stratified by patient age

	No. of Deaths Among Recipients/No of Recipients		HR (95% CI)	*p*-value
	Male Donors (Reference)	Ever-Pregnant Female Donors		
No-Mixture Cohort				
Recipient age, y				
0-17	107/2251	124/2556	1.634 (1.023 to 2.610)	0.040
18-50	84/1170	94/1296	1.501 (0.979 to 2.303)	0.063
51-70	598/4292	645/4775	1.103 (0.913 to 1.332)	0.309
≥71	933/4499	1010/5042	1.064 (0.900 to 1.257)	0.468
p-value for interaction				0.09500
Single-Transfusion Cohort				
Recipient age, y				
0-17	53/1993	70/2287	2.839 (1.576 to 5.117)	0.001
18-50	16/411	23/504	2.291 (0.885 to 5.928)	0.088
51-70	129/1686	152/2049	0.794 (0.504 to 1.252)	0.321
≥71	236/2099	287/2539	1.063 (0.775 to 1.457)	0.706
p-value for interaction				0.00031
Full Cohort				
Recipient age, y				
0-17	124/2421	141/2645	1.178 (0.824 to 1.685)	0.368
18-50	146/1565	156/1649	1.432 (1.126 to 1.823)	0.003
51-70	922/5570	969/5917	1.010 (0.911 to 1.120)	0.850
≥71	1346/5748	1423/6127	1.023 (0.932 to 1.123)	0.631
p-value for interaction				0.00036

Numbers represent hazard ratios per transfused unit (95% confidence interval) compared to receiving a unit from a male blood donor. *Reference: male blood donors. †The p-value for interaction is for the trend in interaction across the four presented categories of patient age. All models are adjusted for calendar year, blood group (ABO-RhD), hospital, age of the donor, cumulative number of transfusions and an interaction term for hospital and cumulative number of transfusions. No-mixture cohort: consists of all the follow-up time during which patients either received all their red blood cells exclusively from male donors or while they received all their red blood cells exclusively from female donors without a history of pregnancy (never-pregnant donors), or exclusively from female donors with a history of pregnancy (ever-pregnant donors). Single transfusion cohort: consists of patients who received only a single red blood cell transfusion during the time period they were followed.

REFERENCES

1. Pfuntner A, Wier LM, Stocks C. Most Frequent Procedures Performed in U.S. Hospitals, 2011: Statistical Brief #165. In: *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs.* Rockville (MD): Agency for Healthcare Research and Quality (US); 2006.
2. Middelburg RA, Briet E, van der Bom JG. Mortality after transfusions, relation to donor sex. *Vox sanguinis.* 2011;101(3):221-229.
3. Barty RL, Cook RJ, Lui Y, Acker JP, Eikelboom JW, Heddle NM. Exploratory analysis of the association between donor sex and in hospital mortality in transfusion recipients. Transfusion. AABB Annual Meeting 2015 Anaheim; 9/2015, 2015.
4. Bjursten H, Dardashti A, Bjork J, Wierup P, Algotsson L, Ederoth P. Transfusion of sex-mismatched and non-leukocyte-depleted red blood cells in cardiac surgery increases mortality. *The Journal of thoracic and cardiovascular surgery.* 2016;152(1):223-232 e221.
5. Desmarets M, Bardiaux L, Benzenine E, et al. Effect of storage time and donor sex of transfused red blood cells on 1-year survival in patients undergoing cardiac surgery: an observational study. *Transfusion.* 2016;56(5):1213-1222.
6. Chasse M, Tinmouth A, English SW, et al. Association of Blood Donor Age and Sex With Recipient Survival After Red Blood Cell Transfusion. *JAMA internal medicine.* 2016;176(9):1307-1314.
7. Chasse M, McIntyre L, English SW, et al. Effect of Blood Donor Characteristics on Transfusion Outcomes: A Systematic Review and Meta-Analysis. *Transfusion medicine reviews.* 2016;30(2):69-80.
8. Toy P, Popovsky MA, Abraham E, et al. Transfusion-related acute lung injury: definition and review. *Critical care medicine.* 2005;33(4):721-726.
9. West FB, Silliman CC. Transfusion-related acute lung injury: advances in understanding the role of proinflammatory mediators in its genesis. *Expert review of hematology.* 2013;6(3):265-276.
10. Middelburg RA, Van Stein D, Zupanska B, et al. Female donors and transfusion-related acute lung injury: A case-referent study from the International TRALI Unisex Research Group. *Transfusion.* 2010;50(11):2447-2454.
11. Middelburg RA, van Stein D, Atsma F, et al. Alloexposed blood donors and transfusion-related acute lung injury: a case-referent study. *Transfusion.* 2011;51(10):2111-2117.
12. Middelburg RA, Porcelijn L, Lardy N, Briet E, Vrielink H. Prevalence of leucocyte antibodies in the Dutch donor population. *Vox sanguinis.* 2011;100(3):327-335.
13. Evers D, Middelburg RA, de HM, et al. Red-blood-cell alloimmunisation in relation to antigens' exposure and their immunogenicity: a cohort study. *Lancet Haematol.* 2016;3(6):e284-e292.
14. Zalpuri S, Zwaginga JJ, van der Bom JG. Risk Factors for Alloimmunisation after red blood Cell Transfusions (R-FACT): a case cohort study. *BMJ open.* 2012;2(3):e001150-e001150.
15. Rubin DB. Inference and missing data. *Biometrika.* 1976;63(3):581-592.
16. *Stata Statistical Software* [computer program]. Version 14.1. College Station, Texas, USA: StataCorp LP; 2015.
17. Bandyopadhyay S, Brittenham GM, Francis RO, Zimring JC, Hod EA, Spitalnik SL. Iron-deficient erythropoiesis in blood donors and red blood cell recovery after transfusion: initial studies with a mouse model. *Blood transfusion = Trasfusione del sangue.* 2017;15(2):158-164.

CHAPTER 3

Age of platelet concentrates and time to the next transfusion

Transfusion; In press.

Camila Caram-Deelder[1;2]; Johanna G. van der Bom[1;2]; Hein Putter[3]; Anja Leyte[4]; Daan van de Kerkhof[5]; Dorothea Evers[1;6]; Erik A. Beckers[7]; Floor Weerkamp[8]; Francisca Hudig[9]; Jaap Jan Zwaginga[1;6]; Jan M. M. Rondeel[10]; Karen M. K. de Vooght[11]; Nathalie C. V. Péquériaux[12]; Otto Visser[13]; Jonathan P. Wallis[14]; Rutger A. Middelburg[1,2].

1. Center for Clinical Transfusion Research, Sanquin Research; 2. Department of Clinical Epidemiology, Leiden University Medical Center; 3. Department of Medical Statistics, Leiden University Medical Center; 4. Departments of Hematology and Clinical Chemistry, Onze Lieve Vrouwe Gasthuis; 5. Department of Clinical Chemistry and Hematology, Catharina Hospital; 6. Department of Immunohaematology and Blood Transfusion, Leiden University Medical Centre; 7. Department of Internal Medicine-Hematology, Maastricht University Medical Center; 8. Department of Clinical Chemistry, Maasstad Ziekenhuis; 9. LabWest, Haga Teaching Hospital; 10. Department of Clinical Chemistry, Isala Zwolle; 11. Department of Clinical Chemistry and Hematology, University Medical Center Utrecht;12. Department of Clinical Chemistry and Hematology, Jeroen Bosch Hospital; 13. Department of Hematology, VU Medical Center; 14. Department of Hematology, Newcastle upon Tyne Hospitals NHS Foundation Trust.

ABSTRACT

Background: Storage time of platelets concentrates has been negatively associated with clinical efficacy outcomes. The aim of this study was, to quantify the association between storage time of platelet concentrates and interval to the next platelet transfusion for different types of platelet components, stored for up to seven days and transfused to transfusion dependent thrombocytopenic hemato-oncology patients.

Methods: From a cohort of patients from 10 major Dutch hospitals, patients were selected whose transfusion patterns were compatible with platelet transfusion dependency due to hemato-oncological disease . Mean time to the next transfusion and mean differences in time to the next transfusion for different storage time categories (i.e. fresh: <4 days, intermediate: 4-5 days, and old: >5 days) were estimated, per component type, using multilevel mixed-effects linear models.

Results: Among a cohort of 29,761 patients who received 140,896 platelet transfusions we selected 4,441 hemato-oncology patients who had received 12,724 platelet transfusions during periods of platelet transfusion dependency. Transfusion of fresh, compared to old, buffy coat-derived platelets in plasma was associated with a delay to the next transfusion of 6.2 hours(95% confidence interval (CI): 4.5 to 8.0). For buffy coat-derived platelets in PAS-B and C this difference was 7.7 hours (CI:2.2 to 13.3) and 3.9 hours (CI:-2.1 to 9.9) while for apheresis platelets in plasma it was only 1.8 hours (CI: -3.5 to 7.1).

Conclusion: Our results indicate that the time to the next transfusion shortens with increasing age of transfused buffy coat-derived platelet concentrates. This association was not observed for apheresis platelets.

INTRODUCTION

The majority of platelet transfusions are given prophylactically to prevent bleeding in hemato-oncological patients who have become thrombocytopenic as a result of disease-related or treatment-induced severe bone marrow suppression.[1] Prophylactic transfusions are routinely prescribed in case of reversible bone marrow failure, while patients have negligible endogenous platelet production, whenever platelet counts drop below $10x10^9$ platelets/L.[2] In this situation, where the indication for the next transfusion depends only on the platelet count, a lower platelet count increment or reduced platelet survival after platelet transfusion will result in a shorter interval to the next transfusion. Consequently a higher cumulative number of transfusions could be needed with all associated risks and costs.

Several studies have reported associations between storage time of platelet concentrates and outcomes. Recently two published meta-analyses showed that storage time plays an important role in the balance between efficacy, safety, and costs.[3,4] Time to the next transfusion, as an outcome, was found to be reported in eight reviewed papers.[5-12] Four of these studies could be meta-analyzed and estimated the interval between platelet transfusions after transfusion of old platelets to be 0.25 days (CI: 0.13 to 0.38) shorter as compared to transfusion of fresh platelets.[3,5-8]

The influence of storage time on platelet recovery and survival could be affected by the type of platelet component transfused. Different production methods and storage solutions may lead to differences in the stability of stored platelets. In addition, while most previous studies reported storage times up to 5 days only, in the Netherlands platelets stored in plasma or in platelet additive solution (PAS) C can be stored for up to seven days.

The aim of this study was, to quantify the association between storage time of platelet concentrates and interval to the next platelet transfusion for different types of platelet components, stored for up to seven days and transfused to transfusion dependent thrombocytopenic hemato-oncology patients.

METHODS

Study design and population

Platelet transfusion data from two nationwide databases specifically developed to study blood transfusions were merged. As previously described in more detail, these databases included: (i) consecutive transfused patients who received their first ever blood component transfusion between May 2005 and September 2015 in one of the six participating centers of the case cohort study "Risk Factors for Alloimmunisation after red blood Cell Transfusions (R-FACT)",[13,14] and (ii) patients who were transfused between November 2009 and January 2016 in one of the seven participating centers of the Dutch Transfusion Data warehouse (DTD) project.[15] Information on individual components was provided by the national Dutch blood supply (Sanquin bloedbank) and linked to hospital data using the components identification codes. Figure 1 shows the dataflow through the analyses.

The two databases contain similar information about patients and transfusions. The DTD database has additional information about patients' admissions and diagnoses registered via the DBC system.[16] The DBC system is a

Diagnosis Related Group like system for the registration and reimbursement of treatments provided by medical specialists and hospitals. Table 1s (online supplemental material) provides a list of hematological DBC codes and their descriptions.

Patient selection

For the current analyses we wanted to use the interval between consecutive platelet transfusions as a proxy for platelet recovery and survival after transfusion. This proxy will only give a valid estimate of the influence of storage time (i.e. independent of patient characteristics) if we select only patients for whom: 1. platelet transfusions were given at set platelet count triggers, 2. recovery and survival were not negatively affected by the clinical condition or refractoriness of the patient, and 3. patients had sufficiently suppressed bone marrow activity to make endogenous platelet production negligible.

An algorithm was therefore developed aiming to select platelet transfusions given to severely thrombocytopenic, thrombocyte transfusion dependent patients, who had received dose-intensive myelo-suppressive therapy and neither produced endogenous platelets nor had an accelerated platelet consumption. Based on clinical experience with this patient group we set up the following selection criteria.

From the first of these transfusions onwards the algorithm selected every platelet transfusion given within six days of the previous transfusion, as long as the interval between the two platelet transfusions was at least two days (i.e. 48 hours, not consecutive days). Platelet transfusions given within an interval of two days (i.e. the same or the next day) were excluded because they are likely to be the result of an unsuccessful platelet transfusion, or patients with increased consumption, or bleeding, that may not have had any correlation to the storage time of the platelet component. Consecutive transfusions after seven days or more were excluded because i) at exactly seven days they likely represent a pre-determined weekly protocol irrespective of platelet counts (e.g. during weekly outpatient clinic visits); and ii) transfusions intervals bigger than seven days there is likely some endogenous production of platelets, as transfused platelets are unlikely to survive that long in the circulation.

Patients could be included in multiple distinct periods of transfusion dependency if the platelet transfusion free interval between these periods was at least 14 days. For examples of patient selection and definitions of transfusion periods, see supplemental material.

Validation of the patient selection

It was pre-defined that the algorithm would be considered optimal if all selected patients were eligible, even if not all eligible patients were selected. Therefore, priority was given to specificity (i.e. no ineligible patients included) for three reasons: 1. not all patients with an appropriate DBC code are actually eligible for this study, since they could also be clinically unstable, refractory to platelet transfusions, or not being exposed to myelo-suppresive agents (i.e. we expect a maximum sensitivity achievable of about 75%);[17-19] 2. we do not expect any bias if we exclude some of the eligible patients; 3. conversely, inclusion of ineligible patients is expected to dilute the influence of storage time on time to the next platelet transfusion, since patient-related factors will then be more important.

Validation was carried out in the DTD database only, since the R-FACT database didn't contain information on diagnoses. However, since this meant diagnoses were missing for logistical reasons (i.e. which hospital transfused a patient, and in which database did this hospital participate), missingness of diagnoses was considered to be missing completely at random.[20] Therefore, no difference in validity of the selection is expected between the two databases and a valid algorithm for one database can validly be applied to the other.

The exclusion of transfusions after an interval of seven or more days was aimed at excluding both patients with endogenous platelet production and out-patients. Similar to the diagnosis we could only validate the exclusion of out-patients for the DTD database.

Although we could not directly validate the selection for the absence accelerated platelet consumption, our selection criteria already select for this (i.e. patients with accelerated consumption are expected to need transfusions with intervals of less than two days). Therefore, an additional check was unnecessary.

Furthermore, the results were stratified by hospital and patient's age categories to provide insight into the consistency of the algorithm's performance across levels of these variables.

Blood components

Platelets components in the Netherlands are obtained from apheresis or whole blood donations. Whole blood donations are separated into components and the buffy-coats of five donations with preferable identical (always compatible) ABO and Rh D blood group are pooled and stored in plasma or platelet addictive solution (PAS). In the Netherlands, and consequently in our cohort, PAS-B was used until December 2012 and PAS-C from January 2013 onwards.[21,22] Platelets stored in PAS-B had a maximum shelf-life of five days, platelets stored in PAS-C or plasma have a maximum shelf-life of seven days. Further, platelets in plasma can be hyperconcentrated (i.e. plasma removed), by indication, before being transfused. Hyperconcentration is only applied to components stored for five days or less.[23]

Single donor apheresis platelets are drawn by use of apheresis machines and stored in plasma for up to seven days. In the Netherlands the indications for apheresis platelets are the need to transfuse HLA or HPA typed platelets, neonates and adults in special situations (i.e. ABO incompatibility, volume overload, or allergic reactions).[23]

In short, the components analyzed in this paper were (1) apheresis platelets in plasma, (2) apheresis platelets in plasma - hyperconcentrated (3) buffy coat-derived platelets in PAS-B, (4) buffy coat-derived platelets in PAS-C, (5) buffy coat-derived platelets in plasma, and (6) buffycoat-derived platelets in plasma - hyperconcentrated. Patients who received rarely prescribed components (i.e. apheresis platelets in PAS) or who had incorrect or missing data for any of their platelet transfusions were excluded. Storage times were calculated setting the components' donation date as day 0.

Analyses

Relation between storage time and time to next transfusion

Multilevel mixed-effects linear regression models were used. The modes had three nested levels to account for differences between hospitals (i.e. transfusion protocols), multiple transfusion periods per patient, and repeated measurements

within a single transfusion period (e.g. two intervals, in case of three platelet transfusions during one transfusion period). Our outcome of interest was the time to the next platelet transfusion. The determinant of interest was storage time of transfused platelet concentrates. Models were adjusted for confounding variables (day of the week, patient age and sex and blood group compatibility). All variables were included in the model as discrete (i.e. indicators). Compatibility was included in the model as two independent categorical variables: ABO compatibility (identical, minor, major and bidirectional mismatch) and Rh D compatibility (identical, minor and major mismatch). Both variables also included the category "unknown" to indicate when the patient's blood type was unknown. Blood groups of components were all known. The type of blood component (i.e. production method, additional processing and storage solution) was considered a potential effect modifier and therefore not included as a cofounder in the model. Instead results were stratified by component type.

Each platelet transfusion was classified according to the components' storage time on the day of transfusion: '*fresh'* if the transfused component was up to 3 days old, '*intermediate*' if the component was 4 or 5 days old, and '*old*' if the component was 6 or 7 days old.

Predicted means (also known as marginal means, predicted marginal means and predicted marginal distribution) of the time to the next platelet transfusion were derived from the multilevel models to estimate the average predicted outcome and 95% confidence interval for each storage time category.

Sensitivity and exploratory analyses

Several sensitivity analyses were performed to check the robustness of the results. The first one was the "single storage age" analyses as effects of different levels of exposure (in this study mixed storage age) could potentially carry-over across consecutive platelet transfusions. In other words, a poor outcome for the current platelet transfusion could also be the result of the storage age of the previous platelet transfusion.[24,25] To overcome this potential problem, transfusion periods were classified according to their components' storage time: 'only fresh' platelet transfusions if all the transfused components were up to 3 days old, 'only intermediate' platelet transfusions if all their components were 4 or 5 days old and finally 'only old' if all their components were transfused after 6 or 7 days of storage. "Mixed age" were transfusion periods that mix more than one storage time group. Consequently, in the single storage age analyses mixed age transfusion periods were excluded. The second sensitivity analysis was performed by excluding transfusion periods which contained potential outpatient clinic platelet transfusions (i.e. admission and discharge of patients were on the same day) from the analyses. The aim of this exclusion was to rule out that the transfusions in these patients bias the results because the transfusion indication may not be entirely platelet count dependent.

Third, to verify the algorithm performance regarding to diagnoses selection a sensitivity an analysis including only patients with hematological diagnoses was performed.

To further explore possible confounding and effect modification all models were also stratified by storage time in days, patients' sex, and patients' age (dichotomized as <18 years or ≥18 years).

RESULTS

Source population

The two databases (R-fact study and DTD) combined and cleaned included 29,761 patients who received 140,896 platelet transfusions between March 2004 and January 2016 (figure 1). The majority of patients were male (18,260, 61%) and adult (25,502, 86%). They received a median of two platelet transfusions (interquartile range (IQR) 1 to 3). Twenty-one percent (3,638) of the 16,927 patients with diagnoses available had one or more hematological diseases: 1,472 (9%) leukemia, 845 (5%) lymphoma, 663 (4%) myeloma, and 374 (2%) "other hematological diseases". These patients received 47,704 (59%) of all transfusions (Table 1 - *"full cohort"*). Diagnoses were not available to 12,834 patients, 96% of them (12,281) due to lack of information in the source database (R-fact). Only 553 (2%) patients did not have diagnoses available due to missingness. A total of 140,896 platelets units were transfused: 108,823 (77%) buffy coat-derived platelets in plasma, 17,327 (12%) apheresis platelets in plasma, and 14,746 (10%) buffy coat-derived platelets in PAS. ABO and Rh D identical components corresponded to 67% (94,577) and 73% (102,870) of the transfusions. Components were stored on average for 4.0 days (median 4, IQR (3 to 6)). Of all transfused platelets 45,241 (32%) were fresh (<4 days), 57,549 (41%) were of intermediate age (4-5 days) and 38,101 (27%) were old (>5 days). (Table 1 - *'full cohort'*)

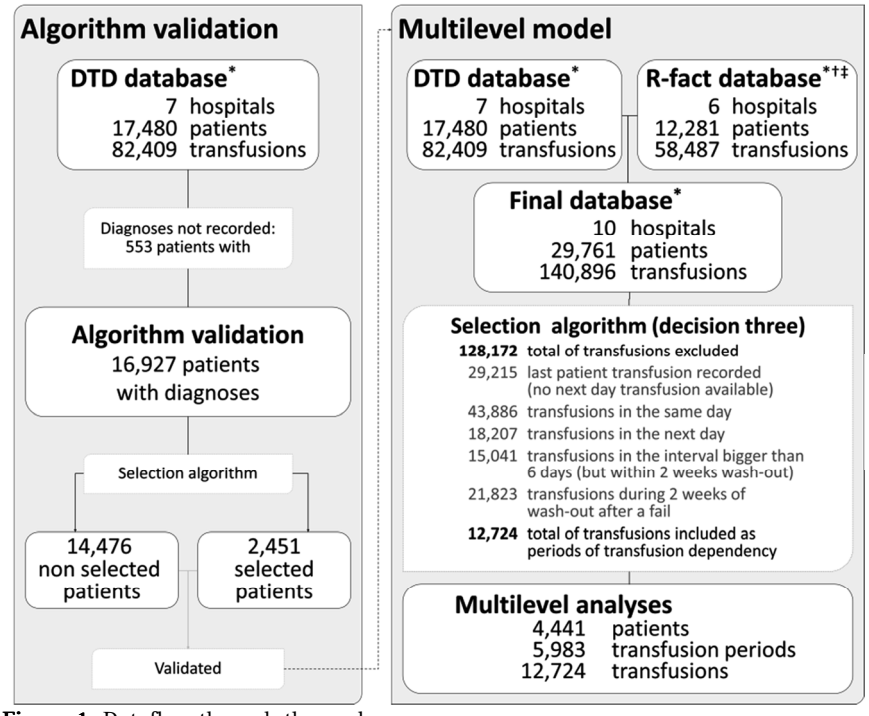


Figure 1: Dataflow through the analyses
DTD: Dutch Transfusion Data warehouse
R-fact: case cohort study "Risk Factors for Alloimmunisation after red blood Cell Transfusions (R-FACT)*Merged to blood supplier database and cleaned: excludes patients who received rarely prescribed products (total of 69 patients) or who had incorrect or missing data (total of 844 patients) †R-fact database does not have diagnoses code. Numbers refer to additional transfusions/patients. Except by hospitals: 6 hospitals in total, 3 new hospitals and 3 hospitals also included in the DTD database‡ three hospitals were common in the DTD and R-fact databases (different follow-up), data duplication was checked by the unique product code

Performance of selection algorithm

Of the 29,761 patients who received platelet transfusions 16,927 had diagnoses available in the source database (i.e. the DTD database), and could be included in the validation of the algorithm (figure 1). 3,638 patients had at least one documented hematological diagnosis and 13,289 did not. Of the 13,289 patients without documented hematological diagnosis 747 were selected by the algorithm while 12,542 were correctly not selected by the algorithm. Thus, the algorithm's overall specificity was 94%. In other words, the probability of not being selected given that the patient does not have any hematological diagnosis was: 12,542/13,289=0.94. From the 3,638 patients with documented hematological diagnoses the algorithm selected 1,704 in one or more periods of transfusion dependency (sensitivity 47%). For children (age <18 years) specificity was 85%, while for adults (≥18 years) specificity was 96%. The algorithm performed similarly for all hospitals (Table 2).

Selected population

Once the algorithm was validated it was applied to the full database. The final selection according to the validated algorithm included 4,441 patients who received 12,724 platelet transfusions in 5,983 transfusion periods (figure 1, table 1). Selected patients received an average of 3.0 transfusions (median 2, IQR: 1 to 3) per transfusion period. 80% of selected patients were adults (median age 56 years, IQR 35 to 65) and 60% male. 1,990 selected patients didn't have diagnoses available. Seventy percent of the 2,451 selected patients, with diagnoses available in the database, had one or more diagnoses of hematological disease. Diagnoses were not available to 1,990 patients, 97% of them (1,940) due to lack of information in the source database (R-fact). Only 50 (1%) patients did not have diagnoses available due to missingness. Leukemia and lymphoma were the most common diagnoses of the selected population (34% and 15%). 78% (9,967) of the transfusions were buffy coat-derived platelets in plasma, 11% (1,442) apheresis platelets in plasma and 10% (1,315) buffy coat-derived platelets in PAS. ABO and Rh D identical components corresponded to 69% (8,733) and 73% (9,334) of the transfusions. 3,649 (29%) of the platelets units were transfused fresh, 5,438 (43%) were transfused at intermedium storage time and 3,637 (29%) units were transfused old. (Table 1 – 'selected cohort')

Table 2: Algorithm performance by patient's age and hospitals

	n	Specificity	Sensitivity
All ages			
All hospitals	16,927	94%	47%
Hospital A	1,505	96%	47%
Hospital B	2,290	96%	47%
Hospital C	4,522	92%	50%
Hospital D	2,201	95%	44%
Hospital E	815	92%	28%
Hospital F	1,868	98%	41%
Hospital G	3,726	93%	51%
Age <18			
All hospitals	2,196	85%	56%
Hospital A	12	NA*	NA*
Hospital B	88	NA*	NA*
Hospital C	877	83%	59%
Hospital D	140	83%	47%
Hospital E	28	NA*	NA*
Hospital F	4	NA*	NA*
Hospital G	1,047	87%	61%
Age ≥18			
All hospitals	14,731	96%	46%
Hospital A	1,493	96%	48%
Hospital B	2,202	97%	48%
Hospital C	3,645	95%	48%
Hospital D	2,061	96%	44%
Hospital E	787	93%	29%
Hospital F	1,864	98%	41%
Hospital G	2,679	96%	48%

*NA: Not available due to the small number of patients

Time to the next transfusion

Figure 2 and table 3 show the time to the next transfusion (in days) for each component and the difference (in hours) between storage time categories. Fresh buffy coat-derived platelets in plasma (<4 days) resulted in a time to the next transfusion of 3.5 days (95% confidence interval (CI): 3.4 to 3.6). Fresh hyperconcentrated buffy coat-derived platelets in plasma resulted in a time to the next transfusion of 3.5 days (CI: 3.3 to 3.6). Fresh buffy coat-derived platelets stored in PAS-C had a time to the next transfusion of 3.1 days (CI: 2.9 to 3.3). Fresh buffy coat-derived platelets stored in PAS-B resulted in a time to the next transfusion of 3.5 days (CI: 3.3 to 3.6). And fresh apheresis platelets in plasma resulted in a time to the next transfusion of 3.3 days (CI: 3.1 to 3.4).

Storage time and time to the next transfusion

Relative to fresh components (<4 days), intermediately stored (4 or 5 days of storage) components had a 3.5 hour shorter (CI: 1.8 to 5.2) interval for buffy coat-derived platelets in plasma, 3.7 hour shorter (CI: -0.6 to 8.0) for hyperconcentrated buffy coat-derived platelets in plasma, 0.1 hour shorter (CI: -5.5 to 5.7) for buffy coat-derived platelets in PAS-C, 7.7 hour shorter (2.2 to 13.3) for buffy coat-derived platelets in PAS-B, 4.7 hour shorter (CI: -0.1 to 9.5) for apheresis platelets in plasma and 0.0 hour longer (CI: -5.6 to -5.7) for hyperconcentrated apheresis platelets in plasma.

Again, relative to fresh components, old components (>5 days) had a 6.2 hours shorter (CI: 4.5 to 8.0) interval for buffy coat-derived platelets in plasma, 3.9 hour shorter (IC: -2.1 to 9.9) for buffy coat-derived platelets in PAS-C, and 1.8 hours shorter (-3.5 to 7.1) for apheresis platelets in plasma.

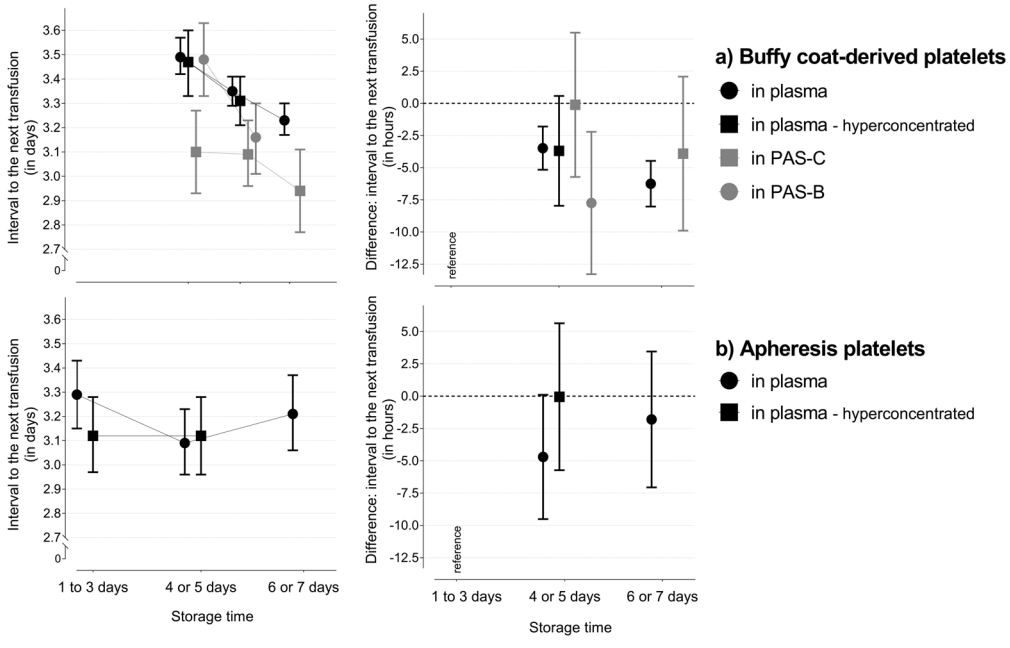


Figure 2 - Interval to the next transfusions (in days) per blood component and difference (in hours)

Table 1: Patient and transfusion characteristics

	Full cohort		Selected cohort	
Patients				
Number of unique patients	29,761		4,441	15%
Transfusion periods	NA		5,983	NA
Female patients	11,062	37%	1,744	39%
Male patients	18,260	61%	2,659	60%
Unknown sex	439	1%	38	1%
Age of patients (in years)*	62	(44-72)	56	(33-65)
<18 years old	4,259	14%	887	20%
≥18 years old	25,502	86%	3,554	80%
Diagnoses per patient				
Not available	12,834	43%	1,990	45%
Not available due to database (R-fact)	12,281	41%	1,940	44%
Not available due to missingness (DTD data)	553	2%	50	1%
Available	16,927	57%	2,451	55%
Others than hematological diseases	13,289	79%	747	30%
Hematological diseases	3,638	21%	1,704	70%
Leukemia†	1,472	9%	844	34%
Chronic leukemia †	238	1%	95	4%
Myeloma†	663	4%	199	8%
Lymphoma†	845	5%	357	15%
Childhood hematological diseases†	204	1%	112	5%
Others hematological diseases†	374	2%	173	7%
Transfusions per patient*	2	(1-3)	2	(1-3)
Transfusions				
Total number of platelets units transfused	140,896		12,724	9%
Buffy coat-derived in plasma	88,802	63%	8,709	68%
Buffy coat-derived in plasma - hyperconcentrated	20,021	14%	1,258	10%
Buffy coat-derived in PAS-C	8,323	6%	625	5%
Buffy coat-derived in PAS-B	6,423	5%	690	5%
Apheresis platelets in plasma	10,966	8%	964	8%
Apheresis platelets in plasma - hyperconcentrated	6,361	5%	478	4%
ABO compatibility‡				
Identical	94,577	67%	8,733	69%
Minor	31,121	22%	2,525	20%
Major	8,249	6%	652	5%
Bidirectional	1,988	1%	154	1%
Unknown	4,961	4%	660	5%
Rh D compatibility‡				
Identical	102,870	73%	9,334	73%
Minor	28,188	20%	2,370	19%
Major	5,581	4%	425	3%
Unknown	4,257	3%	595	5%
Storage time				
1 day	1,999	1%	153	1%
2 days	16,848	12%	1,308	10%
3 days	26,399	19%	2,188	17%
4 days	27,846	20%	2,670	21%
5 days	29,703	21%	2,768	22%
6 days	19,120	14%	1,863	15%
7 days	18,981	13%	1,774	14%
Transfusions per diagnoses				
Not available	59,509	42%	5,747	45%
Not available due to database (R-fact)	58,487	42%	5,677	45%
Not available due to missingness (DTD data)	1,022	1%	70	1%
Available	81,387	58%	6,977	55%
Others than hematological diseases	33,683	41%	1,208	17%
Hematological diseases	47,704	59%	5,769	83%
Leukemia†	24,688	30%	3,594	52%
Chronic leukemia†	3,873	5%	319	5%
Myeloma†	3,696	5%	392	6%
Lymphoma†	5,978	7%	887	13%
Childhood hematological diseases†	3,285	4%	335	5%
Others hematological diseases†	8,710	11%	591	8%

Numbers represent numbers of absolute numbers and percentages unless otherwise indicated.
* median and interquartile range; † multiple answers, percentages of patients with diagnoses available in the database; ‡ Minor compatibility: transfused blood contains antibodies against recipients antigens; Major compatibility: recipients contains antibody against antigens in transfused blood; Bidirectional compatibility: major and minor compatibility combined; Unknown compatibility: patient blood type unknown (blood group of components were all known).

Table 3: Crude and adjusted marginal means and differences of interval to the next transfusions per type of blood component

	Crude			Multilevel (hospital, patient and cycles) adjusted for confounders (weekday, patient age, patient sex, ABO and Rh compatibility)		
	n	Transfusion interval (95% CI) - in days	difference (95% CI) - in hours	n	Transfusion interval (95% CI) - in days	difference (95% CI) - in hours
Buffy coat-derived platelets in plasma						
1 to 3 days	2,099	3.6 (3.5 to 3.6)	reference	2,099	3.5 (3.4 to 3.6)	reference
4 or 5 days	3,435	3.3 (3.3 to 3.3)	-6.8 (-8.3 to -5.2)	3,435	3.3 (3.3 to 3.4)	-3.5 (-5.2 to -1.8)
6 or 7 days	3,175	3.2 (3.2 to 3.3)	-8.9 (-10.5 to -7.3)	3,175	3.2 (3.2 to 3.3)	-6.2 (-8.0 to -4.5)
Buffy coat-derived platelets in plasma - hyperconcentrated						
1 to 3 days	481	3.6 (3.5 to 3.7)	reference	481	3.5 (3.3 to 3.6)	reference
4 or 5 days	777	3.2 (3.1 to 3.3)	-10.3 (-13.7 to -7.0)	777	3.3 (3.2 to 3.4)	-3.7 (-8.0 to 0.6)
6 or 7 days	0	NA	NA	0	NA	NA
Buffy coat-derived platelets in PAS-C						
1 to 3 days	171	3.3 (3.1 to 3.4)	reference	171	3.1 (2.9 to 3.3)	reference
4 or 5 days	276	3.0 (2.9 to 3.2)	-5.6 (-10.8 to -0.4)	276	3.1 (3.0 to 3.2)	-0.1 (-5.7 to 5.5)
6 or 7 days	178	2.9 (2.7 to 3.1)	-8.2 (-13.9 to -2.5)	178	2.9 (2.8 to 3.1)	-3.9 (-9.9 to 2.1)
Buffy coat-derived platelets in PAS-B						
1 to 3 days	338	3.5 (3.3 to 3.6)	reference	338	3.5 (3.3 to 3.6)	reference
4 or 5 days	352	3.1 (3.0 to 3.2)	-8.8 (-13.2 to -4.5)	352	3.2 (3.0 to 3.3)	-7.7 (-13.3 to -2.2)
6 or 7 days	0	NA	NA	0	NA	NA
Apheresis platelets in plasma						
1 to 3 days	321	3.3 (3.1 to 3.4)	reference	321	3.3 (3.1 to 3.4)	reference
4 or 5 days	359	3.0 (2.9 to 3.1)	-6.2 (-10.6 to -1.9)	359	3.1 (3.0 to 3.2)	-4.7 (-9.5 to 0.1)
6 or 7 days	284	3.3 (3.2 to 3.5)	2.0 (-2.6 to 6.6)	284	3.2 (3.1 to 3.4)	-1.8 (-7.1 to 3.5)
Apheresis platelets in plasma - hyperconcentrated						
1 to 3 days	239	3.2 (3.0 to 3.4)	reference	239	3.1 (3.0 to 3.3)	reference
4 or 5 days	239	3.0 (2.9 to 3.2)	-3.6 (-8.9 to 1.6)	239	3.1 (3.0 to 3.3)	0.0 (-5.7 to 5.6)
6 or 7 days	0	NA	NA	0	NA	NA

On average, patients were platelet transfusion dependent for 11.1 days and received platelet transfusions every 3.35 days during that period (total 3.32 platelet transfusions over 11.1 days). When receiving only fresh platelet units, the interval between transfusions would be 3.50 days therefore resulting in a total of 3.17 transfusions compared to an interval of 3.24 days and a total of 3.43 transfusions when receiving only old components. The difference between only fresh and only old would therefore be 0.25 transfusions on average, suggesting that up to one transfusion might be saved on average per 4 patients' admissions or 7% of the patients' transfusions (table 4). Table 4 shows the projected differences for all platelets components.

Sensitivity and exploratory analyses

Results of the different exploratory stratifications and the sensitivity analyses of single storage age, and the analyses after excluding patients without diagnoses available and transfusions in the outpatient clinic were similar to the results presented in the main manuscript (see supplemental material for detailed results).

DISCUSSION

The results of our analyses indicate that the time to the next transfusion decreases as the age of transfused platelet components increases. This decrease was found to be similar, ranging from 0.1 to 7.7 hours, for all buffy-coat-derived platelet components, irrespective of storage solution. Conversely, storage time was not associated with a reduced time to the next transfusion after transfusion of apheresis platelets. The total decrease in the time to next transfusion for buffy-coat derived platelets was a quarter of a day when comparing platelets stored for three days or less to those stored for six or seven days. This difference represents on average 0.25 transfusions per patient's admission.

Although this average of 0.25 less transfusions per admission may not seem to have clinical significance at the individual patient level, since 0.25 units of platelets are never transfused. This figure was estimated at the population level, meaning that some patients will receive one or more units less, while others may not benefit at all.

Table 4: Projected mean difference of total number of transfusions per admission

	Time to the next transfusion (in days)	One transfusion each (days in one admission)*	Difference (days in one admission)
Buffy coat-derived platelets in plasma			
1 to 3 days	3.494	3.18 days	reference
4 or 5 days	3.349	3.31 days	0.14 days
6 or 7 days	3.234	3.43 days	0.26 days
Buffy coat-derived platelets in plasma - hyperconcentrated			
1 to 3 days	3.466	3.20 days	reference
4 or 5 days	3.312	3.35 days	0.15 days
6 or 7 days	NA	NA	NA
Buffy coat-derived platelets in PAS-C			
1 to 3 days	3.100	3.58 days	reference
4 or 5 days	3.095	3.59 days	0.01 days
6 or 7 days	2.937	3.78 days	0.20 days
Buffy coat-derived platelets in PAS-B			
1 to 3 days	3.478	3.19 days	reference
4 or 5 days	3.156	3.52 days	0.33 days
6 or 7 days	NA	NA	NA
Apheresis platelets in plasma			
1 to 3 days	3.289	3.37 days	reference
4 or 5 days	3.093	3.59 days	0.21 days
6 or 7 days	3.214	3.45 days	0.08 days
Apheresis platelets in plasma - hyperconcentrated			
1 to 3 days	3.121	3.56 days	reference
4 or 5 days	3.120	3.56 days	0.00 days
6 or 7 days	NA	NA	NA

* $\left(\frac{\text{time to the next transfusions}}{\text{average length of admission (11.1 days)}}\right)$

NA: not available

Therefore, the positive clinical implications of the increased time between platelet transfusions observed for fresher platelet transfusions are the same as those for a decreased number of transfusions: less acute hemolytic reactions, febrile non-hemolytic reactions, risk of bacterial contamination, transfusion related acute lung injury (TRALI), allergic reactions, and alloimmunization.[26]

Conversely, transfusing only fresh or intermediate aged platelets (i.e. up to five days of storage) would severely affect the outdating and consequently increase the wastage. In the Netherlands it was shown that the outdating decreased from 20% to 10% when the maximum shelf-life of platelets in plasma was increased from five to seven days, corresponding to a preservation of 5,900 components yearly.[22,27]

It is important to realize that a policy of transfusing only fresh platelets to hematological patients would save 7% of these patients' platelet transfusions only when compared to transfusing only old platelets. However, the extended shelf-life of up to seven days does not make all platelets components old, but merely a fraction of them. In our study 27% of the transfused platelets were old (>5 days). Additionally, this gain only applies to hematological stable patients who account for 75% of the platelet transfusions.[17] Thus, the real gain would be a reduction of 1.4% of the total of platelet transfusions given (i.e. 7%×27%×75%) while an extra 10% of all platelet components would be wasted due to out-dating.[27] This results in an increased need for platelet components of about 8.6%.

Our results corroborate recent meta-analyses in which an overall difference in time to the next transfusion between old and new components of 0.25 days (i.e. six hours) for all components combined is reported.[3,4] In the present study the difference in time to the next transfusion between fresh and old platelets varies from 0.2 hours up to 6.2 hours depending on the component type. In the previous meta-analyses, there was no indication of substantial differences between studies investigating buffy coat-derived platelets and studies investigating apheresis platelets. However, the meta-analyses did not include sufficient studies to be able to stratify results per component type, as the present study did. In the current study no association between storage time and time to the next transfusion for apheresis platelets is observed. Besides reflecting a true difference between components this may also be the result of the specific indications for which apheresis platelets are prescribed in the Netherlands (i.e. HLA or HPA typed platelets, neonates, and adults in case of: ABO incompatibility, volume overload, or allergic reactions).[23]

Important strengths of our study are the size of the cohort and the use of a validated algorithm to select the patients of interest. By selecting the patients according to strictly defined transfusion patterns, we included patients whose time to the next transfusion depended on platelet counts. We thereby avoided selecting patients with pre-determined transfusion schedules and patients with insufficient response to platelet transfusions (refractory patients). Our algorithm had excellent performance (high specificity) for the overall population and also for each hospital studied. Patients selected by the algorithm with others than hematological diseases only received general diagnoses codes, like "care trajectory" or "inter-collegial consultation". These patients are potentially (and likely) hemato-oncological patients, who were transfused before a definitive diagnosis was made and recorded in the diagnosis system and consequently in our study database.

A potential limitation of this study is that we did not have information about the hour of the day at which donations and transfusions occurred.

Thus, estimation of the storage time and transfusion interval was only possible in whole days and therefore imprecise. On average, however, donations and transfusions occur mostly in the same time of the day. As a consequence the storage time and interval between transfusions tend to be, most of the time, not far from the presented results.

A seemingly limiting aspect of our study was the lack of diagnoses recorded in one of the databases. However, our sensitivity analyses of only patients with the diagnoses available showed results almost identical to those obtained from the full cohort. We are therefore confident that our algorithm selected the correct patient population allowing us to increase our sample size from 16,927 patients with diagnoses to 29,761 patients in the final database.In conclusion the present study showed that the transfusions interval decreases as the age of transfused platelet components increases, which seems similar for all buffy coat-derived platelet components and irrespective of storage solution. We also show that this decrease is unlikely to outweigh the benefit of reduced outdating and wastage, known to be associated with extended storage times. Furthermore, no decrease in transfusion interval was observed for apheresis platelets, which in the Netherlands are only prescribed for specific indications.

SUPPLEMENTAL MATERIAL

Available at: https://goo.gl/uDPNvD

- Predictive marginal per blood components and patients sex and age
- Examples of selection and period definitions
- DBC hematological codes and descriptions

Sensitivity analyses:

- Single storage age transfusion periods
- No outpatient clinic patients
- Age stratification
- Only patients with hematological diagnoses

ACKNOWLEDGEMENTS

We thank Aad Pors (Sanquin Leiden) and Peter F. Kemper (Dutch Transfusion Data warehouse) for delivering the data.

REFERENCES

1. Fasano RM, Josephson CD. Platelet transfusion goals in oncology patients. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program.* 2015;2015:462-470.
2. Estcourt LJ, Birchall J, Allard S, et al. Guidelines for the use of platelet transfusions. *British journal of haematology.* Feb 2017;176(3):365-394.
3. Kreuger AL, Caram-Deelder C, Jacobse J, Kerkhoffs JL, van der Bom JG, Middelburg RA. Effect of storage time of platelet products on clinical outcomes after transfusion: a systematic review and meta-analyses. *Vox Sang.* May 2017;112(4):291-300.
4. Caram-Deelder C, Kreuger AL, Jacobse J, van der Bom JG, Middelburg RA. Effect of platelet storage time on platelet measurements: a systematic review and meta-analyses. *Vox Sang.* Nov 2016;111(4):374-382.
5. Leach MF, AuBuchon JP. Effect of storage time on clinical efficacy of single-donor platelet units. *Transfusion.* 1993;33(8):661-664.
6. Akkok CA, Brinch L, Lauritzsen GF, Solheim BG, Kjeldsen-Kragh J. Clinical effect of buffy-coat vs. apheresis platelet concentrates in patients with severe thrombocytopenia after intensive chemotherapy. *Vox Sang.* 2007;93(1):42-48.
7. MacLennan S, Harding K, Llewelyn C, et al. A randomized noninferiority crossover trial of corrected count increments and bleeding in thrombocytopenic hematology patients receiving 2- to 5-versus 6- or 7-day-stored platelets. *Transfusion.* 2015.
8. Diedrich B, Ringden O, Watz E, Shanwell A. A randomized study of buffy coat platelets in platelet additive solution stored 1-5 versus 6-7 days prior to prophylactic transfusion of allogeneic haematopoietic progenitor cell transplant recipients. *Vox Sang.* 2009;97(3):254-259.
9. Norol F, Kuentz M, Cordonnier C, et al. Influence of clinical status on the efficiency of stored platelet transfusion. *Br J Haematol.* 1994;86(1):125-129.
10. Benjamin RJ, Goodnough LT, Lopez PI, et al. Fresh (1-2 day-old) vs. aged (4-5 day-old) INTERCEPT platelets and conventional platelets provide comparable count increments. However fresh platelets result in superior hemostasis: results of the SPRINT trial. *Transfusion.* 2003;43:9A.
11. Slichter SJ, Davis K, Enright H, et al. Factors affecting posttransfusion platelet increments, platelet refractoriness, and platelet transfusion intervals in thrombocytopenic patients. *Blood.* 2005;105(10):4106-4114.
12. Heuft HG, Goudeva L, Krauter J, Peest D, Buchholz S, Tiede A. Effects of platelet concentrate storage time reduction in patients after blood stem cell transplantation. *Vox Sang.* Jul 2013;105(1):18-27.
13. Evers D, Middelburg RA, de Haas M, et al. Red-blood-cell alloimmunisation in relation to antigens' exposure and their immunogenicity: a cohort study. *The Lancet. Haematology.* Jun 2016;3(6):e284-292.
14. Zalpuri S, Zwaginga JJ, van der Bom JG. Risk Factors for Alloimmunisation after red blood Cell Transfusions (R-FACT): a case cohort study. *BMJ Open.* 2012;2(3).
15. van Hoeven LR, Hooftman BH, Janssen MP, et al. Protocol for a national blood transfusion data warehouse from donor to recipient. *BMJ open.* Aug 04 2016;6(8):e010962.
16. Westerdijk M, Zuurbier J, Ludwig M, Prins S. Defining care products to finance health care in the Netherlands. *The European journal of health economics : HEPAC : health economics in prevention and care.* Apr 2012;13(2):203-221.
17. Squires JE. Indications for platelet transfusion in patients with thrombocytopenia. *Blood transfusion = Trasfusione del sangue.* Apr 2015;13(2):221-226.
18. Doughty HA, Murphy MF, Metcalfe P, Rohatiner AZ, Lister TA, Waters AH. Relative importance of immune and non-immune causes of platelet refractoriness. *Vox Sang.* 1994;66(3):200-205.
19. Legler TJ, Fischer I, Dittmann J, et al. Frequency and causes of refractoriness in multiply transfused patients. *Ann Hematol.* Apr 1997;74(4):185-189.
20. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology.* Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
21. Ashford P, Gulliksson H, Georgsen J, Distler P. Standard terminology for platelet additive solutions. *Vox Sang.* May 2010;98(4):577-578.
22. van der Meer PF. PAS or plasma for storage of platelets? A concise review. *Transfusion medicine.* Oct 2016;26(5):339-342.
23. de Vries R, Haas F. English translation of the dutch blood transfusion guideline 2011. *Clinical chemistry.* Aug 2012;58(8):1266-1267.
24. Middelburg RA, Briet E, van der Bom JG. Mortality after transfusions, relation to donor sex. *Vox Sang.* Oct 2011;101(3):221-229.
25. Middelburg RA, le CS, Briet E, Vandenbroucke JP, van der Bom JG. A solution to the problem of studying blood donor-related risk factors when patients have received multiple transfusions. *Transfusion.* 9/2010 2010;50(9):1959-1966.
26. Garraud O, Cognasse F, Tissot JD, et al. Improving platelet transfusion safety: biomedical and technical considerations. *Blood transfusion = Trasfusione del sangue.* Mar 2016;14(2):109-122.
27. van der Meer PF. Adverse effects of 'old' versus 'young' blood: also true for platelet concentrates? *Clinical laboratory.* 2011;57(3-4):260-262.

CHAPTER 4

Effect of platelet storage time on platelet measurements: a systematic review and meta-analyses

Vox Sanguinis. 2016; 111: 374–382.

Camila Caram-Deelder[1,2]; Aukje L Kreuger[1,2]; Justin Jacobse[1,2]; Johanna G van der Bom[1,2]; Rutger A Middelburg[1,2].

1. Center for Clinical Transfusion Research, Sanquin Research; 2. Department of Clinical Epidemiology, Leiden University Medical Center.

ABSTRACT

Background: The storage time of platelet products negatively affects bacterial safety and platelet function. However, low maximum storage time increases outdating of valuable products. Thus, to quantify the effect of platelet storage time on platelets measurements after platelet transfusion a systematic review and meta-analyses were performed.

Methods: Reports and meeting abstracts of randomized trials and observational studies, performed in humans, reporting platelets measurements after transfusion of platelet products of different storage times were selected until February 2016. Meta-analyses were performed for four different storage time contrasts, each answering a different question. Random effects models were used to account for substantial heterogeneity and the weighted mean differences calculated.

Results: Our search strategy yielded 4234 studies of which 46 papers satisfied the inclusion criteria. As judged by the 1 hour corrected count increment, transfusions of fresher platelets compared to stored platelets showed better increment. The weighed mean difference varied from 2.11 (95%CI: 1.51 to 2.71) to 2.68 (95%CI: 1.92 to 3.45). For the 24 hour corrected count increment the weighted mean difference varied from 1.36 (95%CI: 0.12 to 2.60) to 1.68 (95%CI: 1.07 to 2.28) depending on the contrast. Recovery and survival of old platelets as percentage of fresh platelets were 81% and 73% for the original definition contrast. For the extended storage contrast recovery and survival were 75% and 68%.

Conclusions: Fresh platelets were superior to old platelets for all platelets measurements and for all storage time contrasts meta-analysed.

INTRODUCTION

Many papers have been published relating storage time of blood products to clinical outcomes and measurements. However, most of these focus on red blood cells.[1-9] Platelets are essential for hemostasis. Patients with thrombocytopenia or thrombocytopathy, due to hematologic malignancies, other blood disorders, bleeding, or medication, require platelet transfusions to prevent or treat bleeding.[6,7]

The storage time of platelet products negatively affects bacterial safety and platelet function.[8,9] However, low maximum storage time increases outdating of valuable products. The balance between avoiding wastage and maintaining product safety and quality determines optimal storage time.[10] Maximum storage of platelets can be three to seven days, depending on the local or national guidelines and the type of product. For example, maximum storage time is three days in Japan[11], four days in Germany[12] and five days in the United States[13] and Brazil[14]. In The Netherlands platelet products can be stored for a maximum of seven days.[15] As blood banks world-wide seek to increase maximum storage times, seven day storage will become more common. The effect seven day storage has on product quality and safety will therefore become ever more important. In 2014 the Food and Drug Administration issued a draft guidance on safety testing and, during their 2015 annual meeting, The American Association of Blood Banks hosted a dedicated session "Paving the Way Towards Implementation of 7 Day Platelets".

Several studies have investigated the effect of storage time of platelets on platelets measurements and other outcomes.[16,17] However, no comprehensive systematic summary and quantification (meta-analyses) of the available evidence has been made to date. The objective of this systematic review and meta-analyses was to quantify the effect of platelet storage time on platelets measurements after platelet transfusion.

METHODS

Search strategy

As pre-specified in the study protocol (online appendix 1), we performed a systematic review to identify all randomized clinical trials and observational studies reporting storage time of platelets products. Potentially relevant papers and meeting abstracts were identified using MEDLINE (PubMed), EMBASE, Cochrane, CINAHL, Academic Search Premier, ScienceDirect and Web of Science databases until February 2016. No restriction on study design, language or year of publication was used (online appendix 2). Non-English papers were translated by native (Chines and German) or fluent (Russian) speakers.

Study selection

Two reviewers independently reviewed, titles and abstracts to select studies reporting platelets storage time and platelets measurements. Pre-specified inclusion criteria were: (i) human: papers reporting exclusively animal studies were excluded; (ii) platelet product transfusion: papers that were exclusively about other blood products or about endogenously produced platelets were excluded; (iii) clinical (performed in patients or volunteers): in vitro, ex vivo, laboratory experiments, and simulation studies were excluded; (iv) storage time: reported as a variable in the paper; (v) original: letters, comments, and reviews not containing any original data were excluded; (vi) platelets measurements: papers that reported at least one of the five platelets measurements (count increment [$\times 10^9$/L]: pre-transfusion platelet count subtracted from post-transfusion platelet

count;[16] corrected count increment [/dm]: count increment corrected for body surface area and platelet product dose;[16] recovery: proportion of platelets recovered from the circulation;[17] survival: mean residual life span;[17] and half-life) and (vii) data necessary for meta-analyses reported: point estimate (i.e. mean or median) and measure of precision (i.e. standard deviation, standard error, interquartile range or range).

Disagreements between reviewers were discussed with a third reviewer. Papers were included for full text assessment if no decision was possible on title and abstract alone. Full text papers were reviewed again for all inclusion criteria. Papers were excluded if the data presented were the same (totally or partially) as those presented in another selected paper. In this case papers were preferred over meeting abstracts and chronologically newer papers were preferred over older ones.

Risk of bias assessment

The risk of bias was evaluated, using "The Cochrane Collaboration's tool for assessing risk of bias" to evaluate randomized clinical trials, and the "Fowkes & Fulton tool" to evaluate both randomized clinical trials and observational studies.[18-20] The items in the Fowkes & Fulton tool are appropriate study design, representative study sample, acceptable control group, quality of measurements and outcome, completeness, confounding, which is similar to the ACROBAT NRSI Cochrane tool for assessing non-randomised studies.[21] For the randomized studies there was perfect agreement between the two tools. Papers with high risk of bias in any of the assessed domains of bias were excluded from the final selection.

Storage time definition

For simplicity only the terms *fresh* and *old* are used throughout this paper. The term *fresh* is used to refer to the storage time group stored for a shorter time than its comparator group (in the same paper). Common synonyms for fresh used in the literature include *new* and *young*. The term *old* is used to refer to the storage time group with the longer storage time. Common synonyms for old include *stored* and *aged*.

Storage time comparisons

To answer different questions regarding the effect of storage time of platelets results were meta-analysed in four different ways.[22] If a paper did not report the results in a way compatible with dichotomizing the data according to one of these definitions, that paper was excluded from that particular analysis.

a) Original definition (as reported): Fresh and old were included in the meta-analysis as reported in the paper. If a paper's results were not presented in two groups the results were dichotomized into fresh if stored ≤3 days and old if stored ≥4 days.

b) Maximum storage 5 days (0-2 vs. 3-5): Papers were included that reported results for zero to two days (fresh) and three to five days (old). This analysis provides a clinically relevant answer to the question whether platelets on the "fresh half" of the storage time spectrum are different from those on the "old half", for the very common situation where the maximum storage time is five days.

c) Extreme difference (0-2 vs. 5-7): To examine the effect of extreme differences in storage time only papers were included if they reported results for zero to two days (fresh) and five to seven days old. This analysis provides the strongest contrast and therefore is the most sensitive indication whether any effect exists or not.

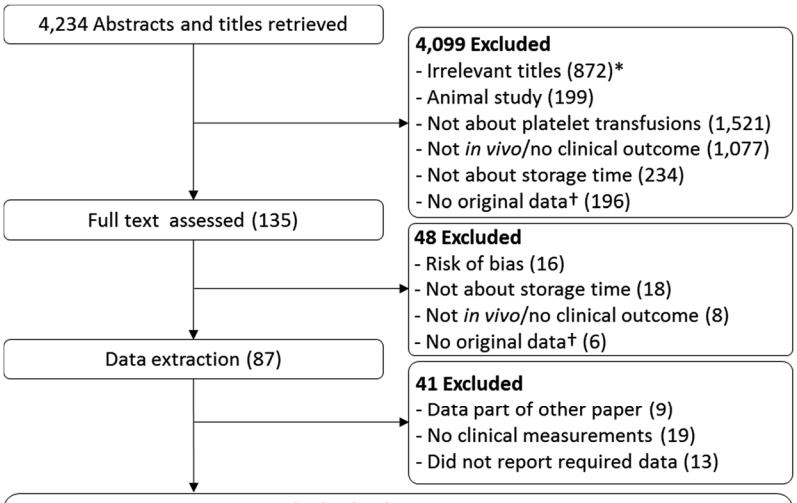
d) Extended storage (0-5 vs. 6-7): In this analysis papers were included that reported results for zero to five days (fresh) and for six or seven days old. It compares "standard maximum

storage" of five days directly to "extended storage" till seven days. It is therefore most relevant to the situation where extended storage is either allowed, or under consideration for implementation.

Each one of these four meta-analyses was performed independently. For all analyses a minimum of five papers (per platelets measurement) was required to estimate the pooled effect. Clinical measurement reported in less than five papers were reported in the selection flowchart (figure 1), but were not included in the meta-analyses. Moreover, for all analyses, results from storage time beyond normal blood banking practice (i.e. >7 days) were disregarded. Pooled effects are presented per platelets measurement.

Data extraction

As specified in the study protocol (online appendix 1), all relevant data reported in the papers were first recorded exactly as reported and subsequently organized and recalculated as described below. Products were grouped into four product groups: apheresis platelets stored in plasma (apheresis plasma), buffy-coat derived platelets stored in plasma (BC plasma), platelet rich plasma (PRP), and buffy-coat derived platelets stored in platelet additive solution (BC PAS). To allow pooling of the data, the original results sometimes needed to be recalculated or transformed:



Final selection: 46 papers

Clinical measurements ‡	Original Definition§ As reported	Maximum storage 5 days§ 0-2 vs. 3-5 days	Extreme difference§ 0-2 vs. 5-7 days	Extended storage§ 0-5 vs. 6-7 days
Total	46 (13)	29 (6)	25 (6)	16 (7)
Corrected count increment	27 (7)	18 (4)	11 (2)	5 (2)
1-h corrected count increment	23 (6)	15 (4)	10 (2)	4 (1)
24-h corrected count increment	23 (7)	15 (4)	8 (2)	4 (2)
Count increment	9 (1)	6 (1)	4 (1)	4 (0)
1-h count increment	6 (1)	4 (1)	3 (1)	2 (0)
24-h count increment	8 (1)	5 (1)	3 (1)	3 (0)
Recovery	18 (6)	10 (2)	13 (4)	9 (5)
Survival	15 (4)	8 (1)	11 (3)	8 (4)
Half-life	4 (1)	3 (1)	3 (1)	1 (1)

Figure 1 – Flowchart Studies Selection

* 886 titles screened (abstracts not available);
† letters/comments/ reviews/reports;
‡ more than one possible outcome per paper
§ between brackets the number "randomised trials"

i. If the standard error of the mean (SEM) was reported, the standard deviation (SD) was calculated: $SD = SEM * \sqrt{n}$;

ii. Mean and standard deviation were calculated from medians, ranges and quartiles[23], since a normal distribution could be expected to be the true underlying distribution from which sampling took place. Only six out of 46 studies did not report their results as normally distributed. We therefore assumed those six were not sufficiently confident of a normal distribution based on their own results alone. Based on the other 40 studies, all sampling from the same underlying distribution, and all reporting a normal distribution, we could be more confident than any individual study;

iii. Similar products (i.e. differences in post-production processing) were merged using standard formulas for combining samples sizes (Σn_i), means ($\sum \overline{x_i} * n_i / \sum n_i$)) and standard deviations ($SD = (\Sigma(n_i - 1)s_i^2 / \Sigma(n_i - 1))^{1/2}$) from multiple groups. Whereas really different products (i.e. different donation procedure or storage medium) presented in the same paper were not merged;

iv. When necessary originally reported categories were merged into the four different definitions of fresh versus old using standard formulas, as described above (item iii);

v. Results presented in hours were recalculated to days;

vi. Platelets measurements reported between zero and four hours after transfusion were considered "1 hour"; platelets measurements reported between eight and 28 hours after transfusion were considered "24 hours".

Analyses

Results were pooled across studies using random effects methods to account for substantial heterogeneity, as indicated by high I^2-values. Weighted mean differences, also known as non-standardized mean differences, were calculated for continuous outcomes. Heterogeneity between studies was assessed using the I^2 statistic. The I^2 value ranges from 0% to 100% and calculates the proportion of variation due to heterogeneity rather than due to chance. Reporting (or publication) bias was analysed using a funnel plot and its asymmetry was assessed using Egger's test.[24] All outcomes (i.e. parameters) were transformed to the same scale to allow the construction of a single funnel plot for all platelets measurements combined. The standardized model was therefore used in this analysis (i.e. as opposed to the non-standardized model used to report the main effects) and all studies were centred around the null effect by subtracting the standardized mean differences per platelets measurement.

Recovery and survival were expressed as percentage recovery and survival achieved with old platelets, compared to fresh platelets. This provides some insight into the order of magnitude of difference to expect, since it allows comparison to the requirements of the Food and Drug Administration (FDA). The FDA requires a minimum of 67% for recovery and 58% for survival, compared to day zero platelets, for any type of platelet product or production process to be allowed into platelets use.[13]

Additional analyses

Additional analyses were performed to clarify whether observed heterogeneity could potentially be attributed to effect modification. Explored possible underlying differences included differences in outcomes, storage times contrasts (analyses a to d), product types, studies populations, and studies design: (i) Funnel plot for each outcome separately; (ii) forest plots for each outcome separately and stratified by different product types and different populations; and (iii) summary mean difference according to whether the study was randomized or not.

RESULTS

Selection

The search retrieved 4234 records. 4099 records were excluded because they were: an exclusively animal study (199); not about platelet transfusions (1521); not in vivo or did not report a platelets outcome (1077); not about storage time (234); did not present original data (196); or because the titles were Irrelevant (872 from the 886 records which abstracts were not available). Upon full text review of the remaining 135 papers a further 48 were excluded because of the above mentioned exclusion criteria (n=32), or because of high risk of bias (n=16, mostly because the fresh and old groups also differed in other respects like storage medium, type of storage bag, storage conditions, type of donation, or production process). Further nine papers were excluded because their data were presented in another selected paper, 19 because they did not report any platelets measurement and 13 because they did not report the data necessary for the meta-analyses. The final selection included 46 papers, 13 randomized trials and 33 observational studies (figure 1). The complete list of selected papers and their qualitative overview can be found in the online supplemental material (appendix 3). Only six papers failed to report normally distributed results. To allow pooling the data their results were recalculated (see methods section for details).

Reported outcomes

Of the 46 selected papers, 27 papers reported corrected count increments (23 reported the 1 hour and 23 reported the 24 hour corrected count increment). Nine papers reported count increment (six papers reported 1 hour and eight reported 24 hour count increment). Eighteen papers reported platelet recovery. Survival was reported in 15 papers and half-life was reported in four (figure 1).

Meta-analyses

Figure 2 shows the funnel plot for all outcomes combined. There is a relative lack of smaller studies (i.e. larger standard error) favouring older platelets, compared to either smaller studies favouring fresh platelets or larger studies. This indicates a bias towards withholding publication of small and therefore statistically unreliable studies showing a benefit of older platelets. Publication bias was present as indicated by Egger's bias coefficient 2.14 (95% confidence interval (CI): 1.59 to 2.70). Half-life did not reach the cut-off of a minimum of five papers and were therefore not included in any of the meta-analyses.

a) Original definition (as reported)

Figure 3a shows the pooled weighted mean differences of fresh platelets minus old platelets. Pooled effect estimates were: 1 hour corrected count increment 2.30 (CI: 1.72 to 2.88); 24 hour corrected count increment 1.68 (CI: 1.07 to 2.28); 1 hour count increment 4.47 (CI: 2.13 to 6.82); 24 hour count increment 4.60 (CI: 0.73 to 8.47); recovery 11.12% (CI: 7.80% to 14.43%), survival 2.08 days (CI: 1.63 to 2.52).

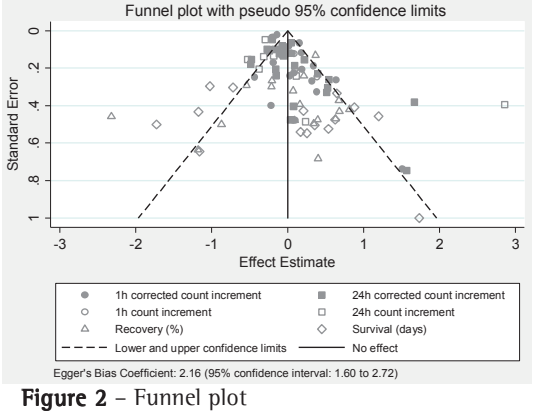


Figure 2 – Funnel plot

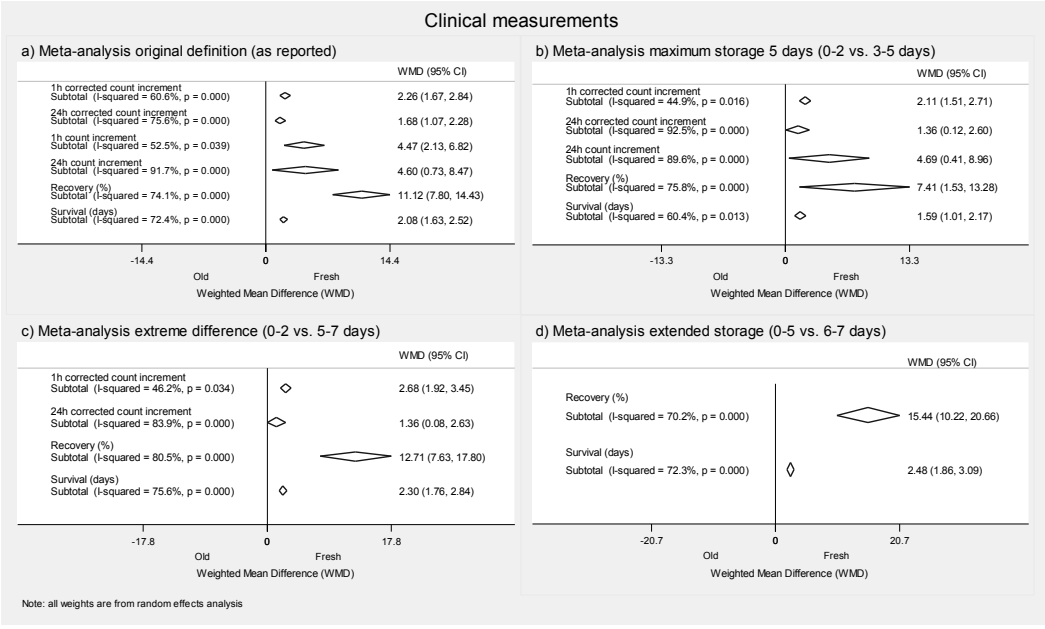


Figure 3 – Summary mean differences between fresh and old platelets products in platelets measurements according to four different definitions of old and fresh

The I^2 ranged from 53% to 92% (table 1 and figure 3a). Based on the pooled means and standard deviation recovery of old platelets was 81% of fresh platelets and survival of old platelets was 73% of fresh platelets (table 1).

b) Maximum storage 5 days (0-2 vs. 3-5 days)

Twenty-nine papers were included in this analysis, 18 papers reported corrected count increment (15 the 1 hour corrected count increment, and 15 the 24 hour corrected count increment) and six reported count increment (four the 1 hour count increment, and five the 24 hour count increment). Recovery and survival were reported in ten and eight papers. The pooled weighted mean differences estimated for fresh minus old were: 1 hour corrected count increment 2.11 (CI: 1.51 to 2.71); 24 hour corrected count increment 1.36 (CI: 0.12 to 2.60); 24 hour count increment 4.69 (CI: 0.41 to 8.96); recovery 7.41% (CI: 1.53% to 13.28%) and survival 1.59 days (CI: 1.01 to 2.17). I^2 ranged from 45% to 90% (table 1 and figure 3b). Recovery and survival of old platelets were 88% and 80% of fresh platelets (table 1).

c) Extreme difference (0-2 vs. 5-7 days)

Twenty-five papers were included in the extreme difference (0-2 vs. 5-7 days) meta-analyses. Ten papers reported corrected count increment as an outcome (11 the 1 hour corrected count increment and eight the 24 hour corrected count increment). Four papers reported count increment (three the 1 hour count increment and three the 24 hour count increment). Recovery, and survival were reported in 13 and 11 papers (figure 1). Figure 3c shows the pooled weighted mean differences for fresh minus old for corrected count increment, recovery and survival.

Count increment did not reach the cut-off of a minimum of five papers. Pooled effect estimates were: 1 hour corrected count increment 2.68 (CI: 1.92 to 3.45); 24 hour corrected count increment 1.36 (CI: 0.08 to 2.63); recovery 12.71% (CI: 7.63% to 17.80%); and survival

2.30 days (CI: 1.76 to 2.84). The I^2 ranged from 46% to 81% (table 1 and figure 3c). Recovery of old platelets was 80% of fresh and survival was 71% (table 1).

d) Extended storage (0-5 vs. 6-7 days)

Sixteen papers compared standard storage (0-5 days) to extended storage (6-7 days). Nine papers reported recovery and eight papers reported survival as an outcome. Corrected count increment, and count increment did not reach the cut-off of a minimum of five papers. The pooled weighted mean differences for fresh minus old were: recovery 15.44% (CI: 10.22% to 20.66%) and survival 2.48 days (CI: 1.86 to 3.09). The I^2 were 70% and 72% (table 1 and figure 3d).Recovery and survival of old platelets were 75% and 68% of fresh platelets (table 1).

Additional analyses

Online supplemental material shows funnel plot for each outcome separately and complete forest plots for each outcome separately, stratified by different product types and different populations. It also presents summary mean difference according to whether the study was randomized or not; and the underlying distribution (absolute numbers) of the weighted mean differences (online appendix 4 and 5). All results were similar to the overall pooled results as presented in the main text, table, and figures.

Heterogeneity, as indicated by I^2 values, was typically much lower in analysed subgroups, especially upon stratification by product type. This indicates product type to be a source of heterogeneity. However, since overall pooled results were very similar to pooled subgroup results, overall results can be used as summary measures. Subgroup results are therefore only reported online appendix 4.

DISCUSSION

Fresher platelets were superior to older platelets for all platelets measurements and all different storage time contrasts investigated.

Strengths of this study include the comprehendsiveness. There were no limitations on the type of outcome, publication date, study design, population and language. Also, search keywords were defined very broadly, including as many papers as possible. The search strategy was applied to many different literature databases and queries for all databases were built by a senior librarian, specialized in performing systematic literature searches. This approach likely ensured that all potentially relevant papers were retrieved.

Table 1 – Mean differences in platelets measurements after transfusion of fresh and old platelets products according to four different definitions of fresh and old

	Original definition as reported	Maximum storage 5 days 0-2 vs. 3-5 days	Extreme difference 0-2 vs. 5-7 days	Extended storage 0-5 vs. 6-7 days
1h corrected count increment	2.30 (1.72 to 2.88)	2.11 (1.51 to 2.71)	2.68 (1.92 to 3.45)	-
24h corrected count increment	1.68 (1.07 to 2.28)	1.36 (0.12 to 2.60)	1.36 (0.08 to 2.63)	-
1h count increment	4.47 (2.13 to 6.82)	-	-	-
24h count increment	4.60 (0.73 to 8.47)	4.69 (0.41 to 8.96)	-	-
Recovery (%) old as % of fresh*	11.12 (7.80 to 14.43) 81%	7.41 (1.53 to 13.28) 88%	12.71 (7.63 to 17.80) 80%	15.44 (10.22 to 20.66) 75%
Survival (days) old as % of fresh*	2.08 (1.63 to 2.52) 73%	1·59 (1.01 to 2.17) 80%	2.30 (1.76 to 2.84) 71%	2.48 (1.86 to 3.09) 68%

Values are weighted mean differences fresh minus old (95% confidence interval) or percentages (%) *old as percentage of fresh

From all selected papers the maximum possible amount of available data were retrieved. Data reported in ways that did not allow pooling (e.g. medians and ranges or interquartile ranges), were recalculated into means and standard deviations, which do allow pooling. Data were extracted from graphs when necessary. Therefore, we were able to pool the results and perform the meta-analyses on data from as many papers as possible.

Another important strength of this study is the quality of included data. Risk of bias was assessed in two different ways and we found perfect agreement between the two assessment tools. Out of 135 studies reporting at least one platelets measurement 16 were excluded based on the risk of bias assessment. Of the remaining studies data that allowed for pooling of results in the meta-analyses could be extracted from 46.

A possible limitation is that not enough randomized trials were included to perform a meta-analysis restricted to randomized trials. However, to have full transparency of our reporting, we showed results stratified between randomized trials and non-randomized trials in the supplemental material. All results in these analyses were in the same direction and in the same magnitude as those presented in the main text.

Another remark to be made is about the high heterogeneity between the studies measured as I^2. As recommended by The Cochrane, besides verifying the data and exploring the heterogeneity, a random-effects meta-analysis was performed.[25]

We found indications of the presence of publication bias. The funnel plot shows a slight preference for smaller studies favouring fresher platelets and Egger's bias coefficient also indicates the presence of publication bias. However, the funnel plot is centred around zero by subtracting the standardized mean effect. Therefore, the largest observed "negative effect", is in reality still an effect in favour of fresher platelets. Thus, although publication bias may have had a minor effect on the size of our effect estimates, it seems unlikely that this could have materially influenced our conclusions.

These potential consequences of transfusing older platelets, however, have to be put in perspective relative to the consequences of supplying exclusively fresher platelets. The Dutch blood supply organization (Sanquin) switched to extended storage of platelets (i.e. maximum storage of seven days instead of five) in 2002. This prolongation of storage time reduced outdating from 20% to about 10%, reducing cost and increasing platelet availability.[26]

In conclusion, our results indicate that fresh platelets are more likely to result in a successful transfusion than old platelets. With successful transfusion defined as a count increment based measurement being above a specific threshold. However, as currently judged by means of a corrected count increment, the success of a transfusion results from a mixing of effects of patient and product related factors. To be clinically relevant the judgment of success of a transfusion should depend on patient related factors only and be separated from product related factors as much as possible. So besides body surface area and platelet dose of the product, storage time should also be taken into account, to arrive at an even better corrected count increment to judge the success of transfusions. We therefore recommend more research into a storage time independent measure for the success of a platelet transfusion.

SUPPLEMENTAL MATERIAL

Available at: https://goo.gl/uz4dzz or http://onlinelibrary.wiley.com/doi/10.1111/vox.12443/full#footer-support-info

- Appendix S1 – Protocol
- Appendix S2 – Search strategy: queries
- Appendix S3 – Reference list and qualitative overview – Included papers
- Appendix S4 – Funnel plot per outcome; forest plots of weighted mean differences per outcome, product group and population (patients/volunteers); and Summary mean differences according to study design (RCT/Non-RCT)
- Appendix S5 – Underlying distributions (absolute numbers) of the weighted mean differences

ACKNOWLEDGEMENTS

We wish to thank Jan Schoones (senior medical librarian at the Walaeus library; LUMC) for his help with developing the search queries, Pieter van der Meer for his input about platelets product, Ruifang Li for the Chinese papers translation, Saurabh Zalpuri for the translation of a Russian paper, Oliver Zilch for the translation of a German paper, and Dirk de Korte, Leo van de Watering and Marian van Kraaij for their useful comments on our manuscript.

REFERENCES

1. Cohen B, Matot I. Aged erythrocytes: a fine wine or sour grapes? *Br J Anaesth.* 2013;111 Suppl 1:i62-70.
2. Flegel WA, Natanson C, Klein HG. Does prolonged storage of red blood cells cause harm? *Br J Haematol.* 2014;165(1):3-16.
3. van de Watering LM. Age of blood: does older blood yield poorer outcomes? *Curr Opin Hematol.* 2013;20(6):526-532.
4. van de Watering LM. Effects of red blood cell storage in heavily transfused patients. *Curr Opin Anaesthesiol.* 2013;26(2):204-207.
5. Zimring JC. Fresh versus old blood: are there differences and do they matter? *Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program.* 2013;2013:651-655.
6. Estcourt LJ. Why has demand for platelet components increased? A review. *Transfus Med.* 2014;24(5):260-268.
7. American Society of Anesthesiologists Task Force on Perioperative Blood T, Adjuvant T. Practice guidelines for perioperative blood transfusion and adjuvant therapies: an updated report by the American Society of Anesthesiologists Task Force on Perioperative Blood Transfusion and Adjuvant Therapies. *Anesthesiology.* 2006;105(1):198-208.
8. Lee CK, Ho PL, Lee KY, et al. Estimation of bacterial risk in extending the shelf life of PLT concentrates from 5 to 7 days. *Transfusion.* 2003;43(8):1047-1052.
9. Seghatchian J, Krailadsiri P. The platelet storage lesion. *Transfus Med Rev.* 1997;11(2):130-144.
10. Veihola M, Aroviita P, Linna M, Sintonen H, Kekomaki R. Variation of platelet production and discard rates in 17 blood centers representing 10 European countries from 2000 to 2002. *Transfusion.* 2006;46(6):991-995.
11. ISBTWEB. Transfusion Today - Quarterly newsletter of the International Society of Blood Transfusion. 03/2007 (70):9-10. http://www.isbtweb.org/fileadmin/user_upload/Transfusion_Today/2007/2007-70.pdf
12. Vollmer T, Schottstedt V, Bux J, Walther-Wenke G, Knabbe C, Dreier J. Bacterial screening of platelet concentrates on day 2 and 3 with flow cytometry: the optimal sampling time point? *Blood transfusion = Trasfusione del sangue.* 2014;12(3):388-395.
13. Slichter SJ, Bolgiano D, Corson J, Jones MK, Christoffel T, Pellham E. Extended storage of autologous apheresis platelets in plasma. *Vox Sang.* 2013;104(4):324-330.
14. ANVISA. Resolução RDC nº 57, de 16 de dezembro de 2010. In: Brazilian Health Surveillance Agency (Agência Nacional de Vigilância Sanitária - ANVISA); 2010.
15. de Vries R, Haas F. English translation of the dutch blood transfusion guideline 2011. *Clinical chemistry.* 2012;58(8):1266-1267.
16. Dijkstra-Tiekstra MJ, Pietersz RN, Hendriks EC, Reesink HW, Huijgens PC. In vivo PLT increments after transfusions of WBC-reduced PLT concentrates stored for up to 7 days. *Transfusion.* 2004;44(3):330-336.
17. Aubuchon JP, Herschel L, Roger J, et al. Comparison of computerized formulae for determination of platelet recovery and survival. *Transfusion.* 2005;45(7):1237-1239.
18. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Bmj.* 2011;343:d5928.
19. Fowkes FG, Fulton PM. Critical appraisal of published research: introductory guidelines. *Bmj.* 1991;302(6785):1136-1140.
20. Pavenski K, Rebulla P, Duquesnoy R, et al. Efficacy of HLA-matched platelet transfusions for patients with hypoproliferative thrombocytopenia: a systematic review. *Transfusion.* 2013;53(10):2230-2242.
21. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI). 2014. www.riskofbias.info. Accessed 03-02-2016.
22. Middelburg RA, Wiersum-Osselton JC, van de Watering LM, van der Bom JG. Observational etiologic research: part 1--The etiologic research question: it requires DATA. *Transfusion.* 2013;53(11):2606-2608.
23. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC medical research methodology.* 2014;14:135.
24. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Bmj.* 1997;315(7109):629-634.
25. Higgins JPT, Green S, (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. *The Cochrane Collaboration* 2011. Available from www.cochrane-handbook.org.
26. van der Meer PF. Adverse effects of 'old' versus 'young' blood: also true for platelet concentrates? *Clinical laboratory.* 2011;57(3-4):260-262.

CHAPTER 5

Effect of storage time of platelet products on clinical outcomes after transfusion: a systematic review and meta-analyses

Vox Sanguinis. 2017; 112: 291–300.

Aukje L. Kreuger[1,2]; Camila Caram-Deelder[1,2];
Justin Jacobse[1,2]; Jean-Louis Kerkhoffs[1,3];
Johanna G. van der Bom[1,2]; Rutger A. Middelburg[1,2].

1. Center for Clinical Transfusion Research, Sanquin Research; 2. Department of Clinical Epidemiology, Leiden University Medical Center; 3. Department of Hematology, Hagaziekenhuis.

ABSTRACT

Background: Prolonged storage improves availability of platelet products, but could also influence safety and efficacy. This systematic review and meta-analyses summarizes and quantifies the evidence of the effect of storage time of transfused platelets on clinical outcomes.

Methods: A systematic search in seven databases was performed up to February 2016. All studies reporting storage time of platelet products and clinical outcomes were included. To quantify heterogeneity, I^2 was calculated, and to assess publication bias, funnel plots were constructed.

Results: Twenty-three studies reported safety outcomes and fifteen efficacy outcomes. The relative risk of a transfusion reaction after old platelets compared to fresh platelets was 1.53 (95% confidence interval (CI): 1.04 to 2.25) (12 studies). This was 2.05 (CI 1.47 to 2.85) before and 1.05 (CI 0.60 to 1.84) after implementation of universal leukoreduction. The relative risk of bleeding was 1.13 (CI 0.97 to 1.32) for old platelets compared to fresh (5 studies). The transfusion interval was 0.25 days (CI: 0.13; 0.38) shorter after transfusion of old platelets (4 studies). Three studies reported use of platelet products, two for hematological patients, one for trauma patients. Selecting only studies in hematological patients, the difference was 4.51 units (CI1.92; 7.11).

Conclusion: Old platelets increase the risk of transfusion reactions in the setting of non-leukoreduction, shorten platelet transfusion intervals, thereby increase the numbers of platelet transfusions in hematological patients, and may increase the risk of bleeding.

INTRODUCTION

Platelets are transfused to prevent or treat bleeding complications in patients with thrombocytopenia or platelet dysfunction.[1] Platelet products can be stored for a maximum of 4-7 days, depending on national guidelines and type of product.[2-5] During the period 2000-2002, a survey found the mean annual discard rate for 17 blood banks in 10 countries to be 13% (range 6.7-25%). As outdating was the main reason for discarding platelet products, prolonging storage is likely to reduce the number of discarded units.[6] However, *in vitro* studies demonstrated a gradual loss of platelet function during storage at room temperature, which is known as the 'storage lesion'.[7]

We previously performed a systematic review and meta-analyses on the effect of storage time at room temperature on clinical measurements. In these meta-analyses, older platelets had inferior results on all endpoints as compared to fresher products.[8] However, the clinical implications of these effects are not clear. [9,10] Therefore, the aim of the current systematic review and meta-analyses is to quantify the effect of storage time of platelet products on clinical outcomes after transfusion.

METHODS

The search strategy, study selection, methods for assessing the risk of bias, and the data extraction were described previously and are in accordance with a pre-specified study protocol.[8]

Search strategy

In brief, a systematic search was applied to seven databases: MEDLINE (Pubmed), EMBASE, Cochrane, CINAHL, Academic Search Premier, ScienceDirect and Web of Science. Results were checked for missing relevant papers by experts in the field and the search strategy was adapted as needed. The search was last updated and performed in February 2016. The search strategy contained synonyms for platelets, fresh, old, and storage time. No limitations were placed on study design, language or year of publication (S1 table 1).

Study selection

As specified in the study protocol, two reviewers independently screened titles and abstracts for relevance. Inclusion criteria were: performed in humans, concerning platelet transfusion, reporting clinical outcomes, reporting different storage times, and reporting original data. Disagreements were discussed with a third reviewer. The risk of bias was scored according to the "Cochrane Collaboration's tool for assessing risk of bias" for randomized controlled trials [11] and "Fowkes & Fulton tool" for randomized controlled trials and observational studies.[12] The items in the Fowkes & Fulton tool are appropriate study design, representative study sample, acceptable control group, quality of measurements and outcome, completeness, confounding, which is similar as in the ACROBAT NRSI Cochrane tool for assessing non-randomised studies.[13] Papers scoring insufficient on one of these items were excluded.

Studies could only be included in the meta-analyses if they reported both a point estimate and a measure of precision. Further, studies needed to report an effect measure which could be recalculated to allow pooling with data from other studies (e.g. some studies reported only mean storage time in cases and controls, whereas risk ratios were reported in other studies). Papers written in other languages than English were translated and data extraction was verified by native speakers.

Data extraction

Storage time, type of outcome, product type, point estimate, and measure of precision were recorded. Authors of included studies were contacted when additional information was needed. If necessary, original results were recalculated in order to enable pooling of the results. In all cases where the underlying distribution could be assumed to be normal, mean and standard deviation were calculated from median, range and quartiles.[14] Results expressed in hours were recalculated to days.

Categorization

Storage time was dichotomised into fresh and old. Where storage time was already dichotomised, the reported dichotomisation was maintained. Most papers defined fresh as ≤3 days and old as ≥4 days. Therefore these definitions were used to summarize results if papers reported multiple storage time categories, using standard formulas for combining samples sizes (Σn_i), means ($\sum \overline{x_i} * n_i / \sum n_i$)) and standard deviations ($SD = (\Sigma(n_i - 1)s_i^2 / sqrt[\Sigma(n_i - 1)]$) from multiple groups. Results were grouped by product: apheresis, pathogen-reduced apheresis (PR_aph), buffy coat in plasma (BC_plasma), buffy coat in platelet additive solution (BC_PAS), pathogen reduced buffy coat in platelet additive solution (PR_BC PAS), and platelet rich plasma (PRP). If papers reported results concerning different products, these were handled as separate results.

Outcomes

Papers reporting laboratory measurements (i.e. corrected count increments, count increment, platelet recovery, survival, half-life) were reported elsewhere.[8]

Outcomes related to safety aspects were categorized into transfusion reactions, as defined by Delaney et al.;[15] complications, including other adverse events; mortality; and length of hospital stay. In-hospital mortality for trauma patients was assumed to be equivalent to 60 day mortality, if no additional data were available. In other words, we assumed that is was very unlikely that trauma patients who were discharged alive subsequently died within 60 days. . The cut-off point of 60 days was chosen, as these data were available in other papers reporting mortality.

Outcomes related to efficacy aspects were categorized into bleeding; transfusion interval; transfusion need (i.e. number of platelet, red blood cell, and plasma transfusions, or amount of cryoprecipitate during hospital stay or period of five days, as reported); repeated transfusion within 24 hours; and haemostatic potential as measured by thromboelastography.

Statistical analyses

For studies reporting only incidences of transfusion reactions, complications, mortality, and bleeding, the relative risk was calculated using standard formulas.[16] The corresponding 95% confidence intervals were calculated using Fisher's exact test. Standard errors were determined from the confidence intervals. For case control studies, odds ratios were calculated with standard errors according to the formula of Woolf.[17] The included case control studies selected controls in a way which allowed the reported odds ratios to be interpreted as relative risks.[18] These odds ratios were therefore treated as relative risks in all analyses. Relative risks reflecting the risk of stoppage of bleeding, or improvement in bleeding rate were recalculated to reflect the risk of no stopping of bleeding or no improvement of bleeding rate.

For continuous outcomes, weighted mean differences (WMD) were calculated. If more than ten studies were included, a pre-specified subgroup analysis was performed, based on product type (i.e. before or after implementation of universal leukoreduction). Metaregression was performed to examine the impact of product type on the pooled estimate. The adjusted R-squared ($R^2_{adj} = (\hat{\tau}^2_0 - \hat{\tau}^2)/\hat{\tau}^2_0$) was calculated to examine the proportion of heterogeneity explained by product type. A sensitivity analysis was performed, excluding the studies with the largest standard errors and meeting abstracts.

To assess the risk of publication bias, funnel plots were generated and Egger's bias coefficient was calculated.[19] A single funnel plot was made for all continuous endpoints combined. To standardize all outcomes to the same scale, the standardized mean difference (SMD) was calculated for each comparison. The standardized mean difference expresses the size of the intervention effect in each comparison relative to the standard deviation estimated in that comparison.[20] All studies were centred around the point of no effect by subtracting the pooled standardized mean difference for each outcome from the standardized mean difference for that outcome of each comparison.

Heterogeneity was quantified by the I^2 statistic.[21] To account for substantial heterogeneity a random effects model was used for all meta-analyses. As a sensitivity analysis, we performed a meta-analysis including only the observational studies. All statistical analyses were performed using Stata version 14, packages metan and metareg.

RESULTS

Selection

The literature search yielded 4,234 papers, of which title and abstract were screened for the predefined inclusion criteria, as described previously.[8] Following selection on inclusion criteria and the risk of bias, 32 studies, reporting 59 unique comparisons, were included in this systematic review (figure 1). This included five meeting abstracts and 27 original papers. Four papers reported on trials in which storage time was randomized. Twenty-three studies reported on observational cohort studies, of which five were secondary analyses on data of randomized trials. Five papers reported on case control studies. Thirty-one papers were written in English and one in Chinese. Included studies are described in more detail in the online supplemental material.

Safety outcomes

Transfusion reactions

One randomized trial, two secondary analyses of randomized trials, nine cohort studies and five case control studies reported transfusion reactions (figure 1). In ten papers different kind of transfusion reactions were reported as one combined endpoint. In three papers transfusion reactions were specified as febrile non-haemolytic transfusion reactions, in two papers as transfusion related acute lung injury (TRALI), in one paper as allergic transfusion reaction, and in one paper as septic transfusion reaction.

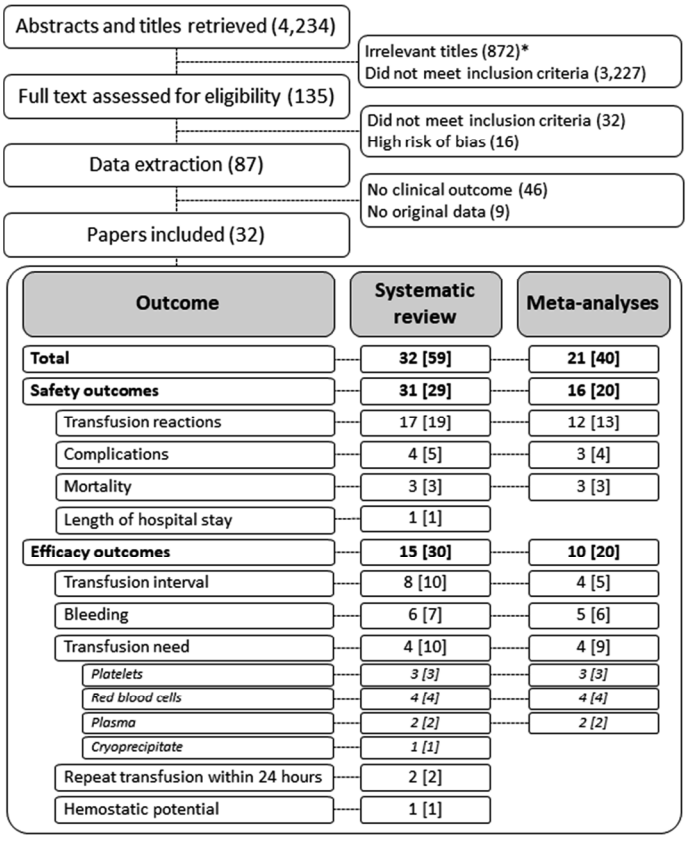


Figure 1 - Flow chart of study selection.

Numbers represent numbers of papers. Some papers reported comparisons for more than one outcome or multiple comparisons for a single outcome. Numbers in square brackets represent the number of unique comparisons.

Twelve studies (thirteen comparisons) were included in the meta-analysis. The pooled risk ratio of old versus fresh platelets was 1.53 (95% confidence interval (CI): 1.04 to 2.25, I^2 83.1%) (figure 2). Before universal leukoreduction was introduced this risk ratio was 2.05 (CI: 1.47 to 2.85, I^2 55.6%) and after introduction it was 1.05 (CI 0.60 to 1.84, I^2 80.8%). The relative risk ratio of leukoreduced products compared to non-leukoreduced products was 0.51 (CI: 0.31 to 0.86, I^2 68.1%). Adjustment for leukoreduction explained 42.36% of heterogeneity. Eggers bias coefficient was 1.62 (p=0.26) (online supplements). Selection of the observational studies yielded a relative risk of 1.05 (CI 0.57 to 1.92) (online supplements). This was similar to the risk ratio in the randomized trial (RR 1.10, CI 0.22 to 5.40). An additional analysis excluding the meeting abstracts and smaller studies, gave similar results (online supplements). Five studies (six comparisons) were excluded from the meta-analysis. three were case control studies comparing mean storage time in both groups, one study did not report the group sizes, and one (two comparisons) only reported a regression coefficient. Of these six comparisons, two reported no difference in incidence of transfusion reactions between both storage time categories in leukoreduced products, three reported an increased incidence after exposure to older non-leukoreduced platelets, and one reported no difference of mean storage time in cases and controls who received leukoreduced as well as non-leukoreduced products (table 1).

Other safety outcomes

Four cohort studies reported complications. Reported complications were: major infection, defined as pneumonia, positive blood culture, leg wound infection, sternal wound infection, or mediastinitis; positive bloodculture; idiopathic pneumonia syndrome; and a composite endpoint of sepsis, ARDS, renal failure, or liver failure. Three studies, four comparisons, were included in the meta-analysis. The pooled risk ratio for these complications of old versus fresh platelets was 1.07 (CI: 0.83; 1.38, I^2 66.6%) (figure 2). One paper could not be included in the meta-analysis, as it reported a hazard ratio of risk of idiopathic pneumonia syndrome, which was 0.84 (CI 0.51 to 1.37).

One randomized trial and two cohort studies reported mortality.[22-24] All were included in the meta-analysis. The pooled risk ratio for mortality was 1.03, (CI: 0.86 to 1.24, I^2 0.0%) (figure 2). The pooled risk ratio in observational studies was 1.03 (CI 0.86 to 1.25) compared to 0.93 (CI 0.29 to 2.96) in the randomized trial was (online supplements).

Length of ICU stay was reported by one study, which found no difference for trauma patients receiving fresh or old platelets.

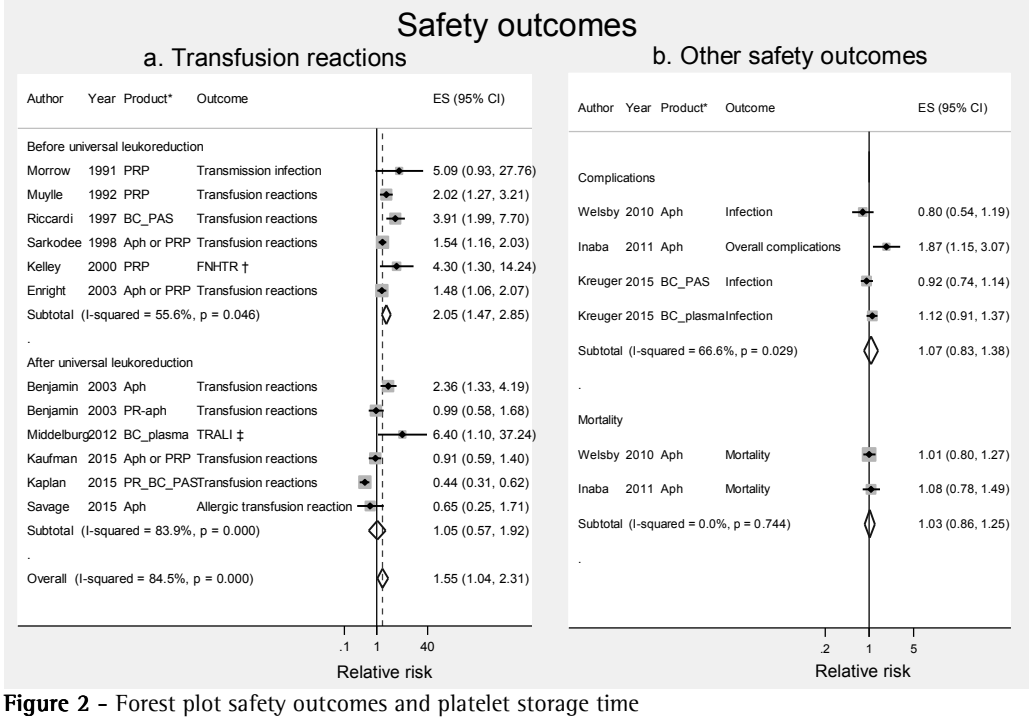


Figure 2 - Forest plot safety outcomes and platelet storage time

a. Meta-analyses of transfusion reactions and platelet storage time, stratified by implementation of universal leukoreduction.
b. Meta-analyses of complications and mortality and platelet storage time. The numbers represent the relative risk of old platelets compared to fresh platelets with corresponding 95% confidence interval for each study.
* Product codes: Aph = apheresis, PRP = platelet rich plasma, BC-PAS = buffy coat stored in PAS, BC-plasma = buffy coat stored in plasma PR = pathogen-reduced
† FNHTR = Febrile non haemolytic transfusion reaction.
‡ TRALI = Transfusion related acute lung injury

Efficacy outcomes

Transfusion interval

Three randomized trials, two secondary analyses of randomized trials and three cohort studies reported a transfusion interval. Four studies (five comparisons) were included in the meta-analysis. The interval between transfusions was 0.25 days (CI: 0.13 to 0.38, I^2 19.5%) longer after transfusion of fresh platelets (figure 3). The weighted mean difference in the observational studies was 0.19 days (CI 0.14 to 0.25) and in the two randomized trials it was 0.42 days (CI 0.10 to 0.75) (online supplements). Four papers (five comparisons) were excluded from the pooled analysis, as these did not provide the necessary measure of precision. Three reported a longer interval following transfusion of fresh platelets. One paper reported no difference in interval following transfusion of apheresis platelet products and a shortened interval after transfusion of fresh pathogen reduced products (table 1). Using the number of transfusions per study as weighing factor, the mean interval reported by the papers excluded from the meta-analysis was 0.14 days.

Bleeding

Two randomized trials, two secondary analyses of randomized trials and two cohort studies papers reported data about bleeding. Reported bleeding endpoints were: incidence of any bleeding symptoms; incidence of bleeding in the central nervous system; percentage of transfusions resulting in lower WHO grade of bleeding; incidence of stopping of gastrointestinal bleeding, haemorrhagic cystitis or epistaxis; proportion of days with bleeding as measured by daily monitoring; and time from transfusion to first bleeding of WHO grade 2. In four studies patients were assessed for bleeding symptoms daily. In two studies medical records were reviewed for bleeding symptoms. Five studies (six comparisons) were included in the meta-analysis. The pooled risk ratio of old platelets versus fresh platelets for any bleeding symptom was 1.13 (CI: 0.97 to 1.32, I^2 38.4%). The pooled risk ratio in observational studies was 1.18 (CI 0.99 to 1.41) and in the two randomized trials the pooled risk ratio was 0.86 (CI 0.58 to 1.27) (online supplements). Exclusion of the meeting abstracts gave similar results (online supplements). One paper could not be included in the pooled analysis, as it reported the time to first ≥WHO grade 2 bleeding (hazard ratio old versus fresh: 1.02 CI: 0.62 to 1.70).

Transfusion need

One randomized trial and three cohort studies reported the need of transfusions. This was reported during hospital stay or during a period of five days. Three papers (three comparisons) were included in the meta-analysis on need of platelet transfusion. The weighted mean difference was 2.76 fewer products (95% CI: -1.11 to 6.64, I^2 84.1%) with fresh platelets compared to old platelets (figure 3). Two studies were performed among hematological patients and one among trauma patients. Selecting only studies in hematological patients yields a weighted mean difference of 4.51 units (CI 1.92; 7.11). The weighted mean difference in the two observational studies was 1.66 units (CI -2.32 to 5.64), and in the randomized trial it was 6.00 units (CI 0.90 to 11.10) (online supplements).

Four papers (four comparisons) were included in the meta-analysis on need of red blood cell transfusions. The weighted mean difference was 0.08 products fewer(95% CI: -0.18 to 0.34, I^2 3.2%) after transfusion of fresh platelets. The weighted mean difference in the observational studies was 0.07 units(CI -0.06 to 0.25), and this

was 2.50 units (CI -1.23 to 6.23) in the randomized trial (online supplements). Two papers (two comparisons) were included in the meta-analysis of need of plasma transfusions. The weighted mean difference was 0.09 products fewer (95% CI: -0.06 to 0.25, I^2 0.0%) after transfusion of fresh platelets (figure 3). One study reported the need of cryoprecipitate, which was not different after transfusion of fresh or old platelets (table 1).

Other efficacy outcomes

One randomized trial and one cohort study reported an increased risk of a repeated transfusion within 24 hours. (table 1). Results from these studies could not be pooled as the storage time of the old platelets in one paper coincided with the storage time of the fresh platelets in the other. One study determined the haemostatic potential of platelets using thromboelastography (TEG) and reported better haemostatic properties of fresh platelets compared to old platelets (table 1).

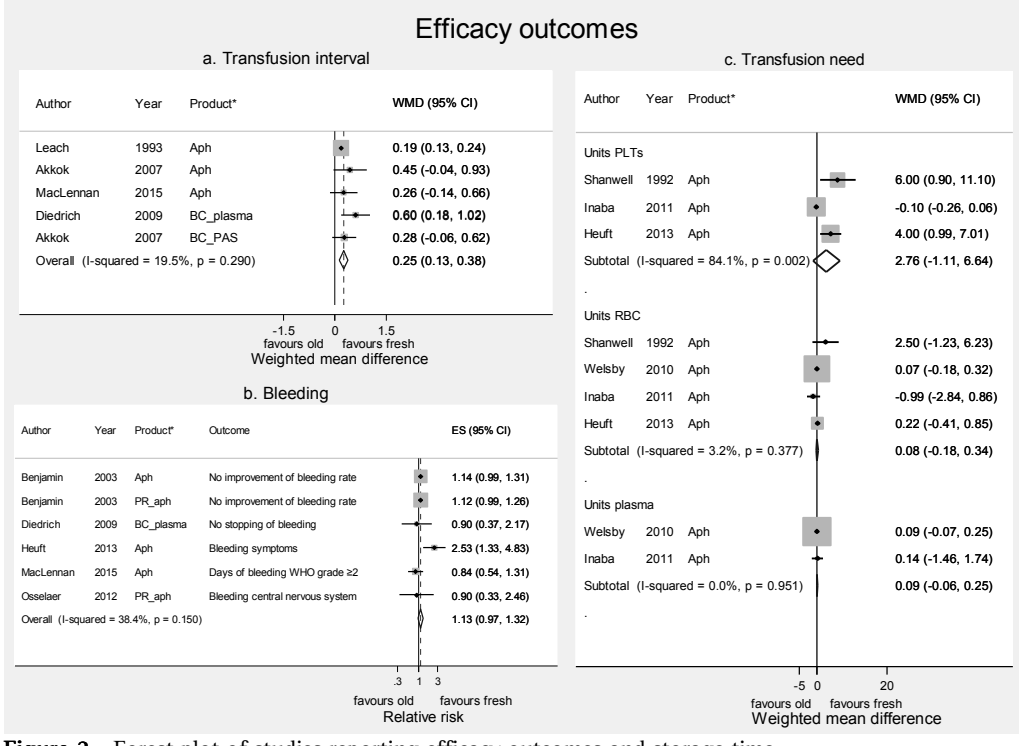


Figure 3 - Forest plot of studies reporting efficacy outcomes and storage time

a. Forest plot of studies comparing the interval between subsequent platelet transfusion in days. The numbers represent the weighted mean difference (WMD), calculated as: 'interval fresh' – 'interval old'.
b. Forest plot of studies reporting the risk of bleeding. The numbers represent the relative risk of old platelets compared to fresh platelets with corresponding 95% confidence interval for each study.
c. Forest plot of studies reporting transfusion need. The numbers represent the weighted mean difference, calculated as 'number of products old' – 'number of products fresh'.
* Product codes: Aph = apheresis, BC-PAS = buffy coat stored in PAS, BC-plasma = buffy coat stored in plasma, PR = pathogen-reduced.
† Results shown for all studies. Selecting only studies in hematological patients yields a weighted mean difference of 4.51 units (CI 1.92; 7.11).

Table 1. Description of studies retrieved by the literature search, but not reporting data necessary for pooling in the meta-analyses

Author and year	Product[a]	Definition fresh	Favours old	No difference	Favours fresh	Definition old	Group size fresh vs. old [*N* transfusions][b]	Outcome (fresh vs. old or controls vs. cases)
Transfusion reactions								
Heddle 1993	PRP or Aph	1–3			■	4–5	Total 65 transfusions	Slope logistic regression $P = 0{,}004$[c]
Lane 1997	Aph	3·1		■		3·8	36 controls vs. 9 cases	3·1 ± 1·1 vs. 3·8 ± 0·7 days[d]
Patterson 2000	PRP_nonL	≤3			■	>3	338 vs. 789 transfusions	Slope linear regression 0·0305; $P < 0{\cdot}001$[c]
Silliman 2003	PRP or Aph	4·2			■	4·5	225 controls vs. 46 cases	4·2 ± 0·1 vs. 4·5 ± 0·2 days, $P = 0{\cdot}014$[d]
Patterson 2000	PRP	≤3		■		>3	306 vs. 1023 transfusions	Slope linear regression 0·009; $P = 0{\cdot}5$[c]
Liu 2013	Aph	2·87		■		2·92	13 controls vs. 16 cases	2·87 ± 0·82 days vs. 2·92 ± 1·03[d]
Complications								
Vande Vusse 2014	PRP or Aph			■			906 patients, 75 cases	HR: 0·84 (CI: 0·51–1·37)[c]
Length of ICU stay								
Inaba 2011	Aph	≤3		■		4–5	128 [205] vs. 253 [380] patients	Median 6 (range 1–181) vs. 6 (1–181) days
Shorter interval between transfusions								
Norol 1994	Aph	≤8 h			■	≤2	141 [141] vs. 141 [141] patients	3·1 vs. 2·3 days
Benjamin 2003	Aph	1–2		■		4–5	697 vs. 1247 transfusions	2·0 vs. 2·0 days; $P = 0{\cdot}97$
Benjamin 2003	PR_aph	1–2	■			4–5	383 vs. 1176 transfusions	1·4 vs. 1·6 days; $P = 0{\cdot}18$
Heuft 2013	Aph	1–4			■	1–5	36 [191] vs. 41 [250] patients	2·0 vs. 1·1 days; $P < 0{,}001$
Slichter 2005	PRP or Aph	<2			■	3–5	Total 5423 transfusions in 525 patients	Difference 0·19 days (CI: 0·12–0·26)
Time to first who ≥2 bleeding								
Triulzi 2012	BC or Aph	3		■		5	156 vs. 217 patients	HR: 1·02 (CI: 0·62–1·70)
Transfusion need: cryoprecipitate								
Inaba 2011	Aph	≤3		■		4–5	128 vs. 253 patients	Median 0 (range 0–33) vs. 0 (0–22) units
Repeated transfusion ≤24 h								
Norol 1994	Aph	≤8 h			■	≤2	88 vs. 88 transfusions	RR: 6·2 (CI: 2·5–15·4)
Duguid 1991	BC_plasma	1–2			■	3–5	77 vs. 40 transfusions	RR: 2·3 (CI: 1·1–4·8)
Haemostatic potential (TEG)[e]								
Roeloffzen 2010	BC_plasma	1–3			■	4–5	35 vs. 35 patients	*K*-time: 27 ± 16 vs. 37 ± 22 min $P = 0{\cdot}03$[e] Alpha angle: 13° ±10° vs. 8°±8° $P = 0{\cdot}02$

Studies can appear more than once if multiple products or multiple end-points were reported. Studies were excluded from the meta-analyses if no measure of precision was reported or if effect measure could not be recalculated in order to allow pooling of results.

■ Conclusion of the paper. 'No difference' means paper found no relevant differences between the groups.

[a] Product codes: aph=apheresis; BC_plasma=buffy coat stored in plasma; PRP=platelet-rich plasma, nonL, non-leucoreduced; PR=pathogen reduced.

[b] Group size is expressed as number of patients [number transfusions], unless otherwise specified.

[c] Storage time analysed as categorical variable, per day.

[d] Storage time in controls vs. storage time in cases.

[e] Thromboelastography measurements: K-time: time until a fixed level of clot firmness is reached in minutes, Alpha angle: rate of clot growth in degrees.

DISCUSSION

To conclude, transfusion of older platelet products was associated with more transfusion reactions before the implementation of universal leukoreduction. This association disappeared after the implementation of universal pre-storage leukoreduction. Transfusion of older platelet products was associated with a shorter time to the next transfusion, a trend toward a higher risk of bleeding, and in hematogical patients an increased need of platelet transfusions. Storage time of platelet concentrates was not associated with the risk of mortality or the consumption of other blood products.

The association between storage time and laboratory measurements (i.e. platelet counts and derivatives thereof) has been reported elsewhere. That study reported inferior results for older platelets for all relevant measurements.[8] The current results suggest that these lower laboratory values are associated with a higher risk of bleeding and a shorter time to the next transfusion. Decreased efficacy of old platelets could explain the increased bleeding risk. Another explanation could be that platelet count is routinely measured on fixed moments, e.g. three times a week. Transfusion of older platelets results in lower increments, leading to a lower platelet count on average in case of a prophylactic transfusion strategy. This could result in an increased bleeding risk.

The increased risk of transfusion reactions in old platelets could be attributed completely to studies performed before the implementation of pre-storage leukoreduction. Leukocytes and leukocyte-derived cytokines are thought to be a major cause of febrile non haemolytic transfusion reactions.[25,26]

With the implementation of universal leukoreduction an absolute risk reduction of 25.1% was expected in the risk of febrile non haemolytic transfusion reactions.[27] The results of the present meta-analyses confirm the beneficial effect of pre-storage leukoreduction on the incidence of transfusion reactions.

An important strength of these meta-analyses is that we were able to pool the available data on bleeding risk. Most studies are powered to study other outcomes and are therefore by themselves inconclusive on bleeding risk. Although different definitions of bleeding are used, we assume storage time has the same effect on all symptoms and it is appropriate to pool the estimates.

Another strength of this study is the broad search strategy. No limits were used for study design, year or language. Therefore, a maximum of available papers reporting clinical effects of storage time have been retrieved and all reported clinical outcomes were studied.

The broad search strategy also returned meeting abstracts, which are possibly more prone to bias. Exclusion of the meeting abstracts did not change the results of the main analyses, indicating these abstracts estimate the same effect. Due to the limited number of randomized trials it was not feasible to perform a sensitivity analysis including only randomized trials. However, the pooled estimates of the observational studies were comparable with the results of the randomized trials. This suggests that the observational studies are reliable, allowing inclusion in the meta-analysis. The relatively large difference between the estimates of the observational studies and the randomized trials in transfusion interval is based on one precise observational study in which the difference in interval was 0.19 days (CI 0.13 to 0.24).

The main limitation of this study is that storage time had to be dichotomized into two broadly defined categories, fresh and old. Most studies reported differences between two groups and defined fresh as storage time of ≤3 days. Therefore it was impossible to compare the safety and efficacy of platelets stored for 1-5 days with platelets stored for 6-7 days. Whereas this is the difference between storage duration used in the Netherlands, compared with other countries.[2-5]

Not all retrieved studies could be included in the meta-analyses, which could potentially induce selection bias. However, the studies excluded from the meta-analysis regarding transfusion interval, reported on average a similar interval as the pooled estimate of the meta-analysis. For the outcomes transfusion reactions and bleeding, the results of the excluded studies pointed in the same direction.

Another limitation of this study is the large heterogeneity between studies reporting transfusion reactions (I^2 83.1%). This is partly due to the difference in effect observed before and after the implementation of universal leukoreduction. Correction for leukoreduction in meta-regression explained 42% of this heterogeneity. Other sources of variation could include the lack of standardized definitions and differences between active and passive monitoring of transfusion reactions. Among studies reporting bleeding symptoms heterogeneity was moderate. This could be due to the fact that several different definitions of bleeding are used and it is measured in different ways. The number of studies reporting on the other outcomes was smaller and therefore it is difficult to detect heterogeneity and publication bias for these outcomes.

In conclusion, the safety and efficacy of platelet products deteriorates during storage. However, leukoreduction reduces the risk of transfusion reactions following transfusion of old platelets effectively. Efficacy of platelet transfusions is reduced after prolonged storage, leading to a shorter interval to the next platelet transfusion. Transfusion of old platelet concentrates might increase the risk of bleeding.

SUPPLEMENTAL MATERIAL

Available at: https://goo.gl/WcsTF9 or http://onlinelibrary.wiley.com/doi/10.1111/vox.12494/abstract#footer-support-info

- Data S1 – Search queries.
- Table S1 – Overview of all included comparisons per product.
- Fig. S1 – Funnel plot of studies comparing incidences of transfusion reactions.
- Fig. S2 – Funnel plot efficacy of platelet transfusion.
- Fig. S3 – Sensitivity analysis: forest plot transfusion reaction, bleeding and platelet storage time.
- Fig. S4 – Sensitivity analysis: forest plot safety outcomes, excluding the randomized trials.
- Fig. S5 – Sensitivity analysis: forest plot efficacy outcomes, excluding the randomized trials.
- Fig. S6 – Forest plot transfusion interval and transfusion need and platelet storage time, standardized analysis.

ACKNOWLEDGEMENTS

We wish to thank the Walaeus/LUMC senior medical librarian Jan Schoones for his efforts with developing the search queries and Ruifang Li for the translation of the Chinese paper.

REFERENCES

1. Estcourt LJ. Why has demand for platelet components increased? A review. *Transfusion medicine (Oxford, England).* Oct 2014;24(5):260-268.
2. Vollmer T, Engemann J, Kleesiek K, Dreier J. Bacterial screening by flow cytometry offers potential for extension of platelet storage: results of 14 months of active surveillance. *Transfus.Med.* 2011;21(3):175-182.
3. ISBTWEB. Transfusion Today - Quarterly newsletter of the International Society of Blood Transfusion. 03/2007 (70):9-10. http://www.isbtweb.org/fileadmin/user_upload/Transfusion_Today/2007/2007-70.pdf
4. Slichter SJ, Bolgiano D, Corson J, Jones MK, Christoffel T, Pellham E. Extended storage of autologous apheresis platelets in plasma. *Vox Sang.* 2013;104(4):324-330.
5. ANVISA. Resolução RDC nº 57, de 16 de dezembro de 2010. Brazilian Health Surveillance Agency (Agência Nacional de Vigilância Sanitária - ANVISA); 2010.
6. Veihola M, Aroviita P, Linna M, Sintonen H, Kekomaki R. Variation of platelet production and discard rates in 17 blood centers representing 10 European countries from 2000 to 2002. *Transfusion.* Jun 2006;46(6):991-995.
7. Seghatchian J, Krailadsiri P. The platelet storage lesion. *Transfusion medicine reviews.* Apr 1997;11(2):130-144.
8. Caram-Deelder C, Kreuger AL, Jacobse J, van der Bom JG, Middelburg RA. The Effect of Platelet Storage Time on Clinical Measurements: a Systematic Review and Meta-analyses. *accepted for publication Vox Sanguinis.*
9. Dijkstra-Tiekstra MJ, Pietersz RN, Huijgens PC. Correlation between the extent of platelet activation in platelet concentrates and in vitro and in vivo parameters. *Vox sanguinis.* Nov 2004;87(4):257-263.
10. Heddle NM, Arnold DM, Webert KE. Time to rethink clinically important outcomes in platelet transfusion trials. *Transfusion.* Feb 2011;51(2):430-434.
11. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Bmj.* 2011;343:d5928.
12. Fowkes FG, Fulton PM. Critical appraisal of published research: introductory guidelines. *Bmj.* May 11 1991;302(6785):1136-1140.
13. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI). 2014. www.riskofbias.info. Accessed 03-02-2016.
14. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC medical research methodology.* 2014;14:135.
15. Delaney M WS, Bercovitz RS, Cid J, Cohn C, Dunbar N, Apelseth TO, Popovsky M, Stanworth SJ, Tinmouth A, Watering L van de, Waters JH, Yazer M, Ziman A, for the Biomedical Axcellence for Safer Transfusion (BEST) Collaborative. Prevention, Diagnosis, and Treatment of Transfusion Reactions: Evidence-Based Review & Clinical Guideline. *Lancet (London, England).* 2016 Dec 3 2016;388(10061):2825-2836.
16. Clopper CJP, E.S.;. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika.* 1934;26:404-413.
17. Woolf B. On estimating the relation between blood group and disease. *Annals of human genetics.* Jun 1955;19(4):251-253.
18. Miettinen O. Estimability and estimation in case-referent studies. *American journal of epidemiology.* Feb 1976;103(2):226-235.
19. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Bmj.* Sep 13 1997;315(7109):629-634.
20. Higgins JPT GS. Cochrane Handbook for Systematic Revieuws of Interventions. *The cochrane collaboration.* 2011;Version 5.1.0.
21. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ (Clinical research ed.).* Sep 6 2003;327(7414):557-560.
22. Inaba K, Branco BC, Rhee P, et al. Impact of the duration of platelet storage in critically ill trauma patients. *J Trauma.* 2011;71(6):1766-1773.
23. Welsby IJ, Lockhart E, Phillips-Bute B, et al. Storage age of transfused platelets and outcomes after cardiac surgery. *Transfusion.* 2010;50(11):2311-2317.
24. Shanwell A, Larsson S, Aschan J, Ringden O. A randomized trial comparing the use of fresh and stored platelets in the treatment of bone marrow transplant recipients. *Eur.J Haematol.* 1992;49(2):77-81.
25. Heddle NM. Pathophysiology of febrile nonhemolytic transfusion reactions. *Current opinion in hematology.* 1999;6(6):420-426.
26. Slichter SJ. Evidence-based platelet transfusion guidelines. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program.* 2007:172-178.
27. Vamvakas EC, Blajchman MA. Universal WBC reduction: the case for and against. *Transfusion.* May 2001;41(5):691-712.

CHAPTER 6

Storage time of platelet concentrates and the diagnosis of platelet refractoriness

Submitted

Camila Caram-Deelder[1,2]; Johanna G. van der Bom[1,2]; Jean-Louis H. Kerkhoffs[1,3]; Rutger A. Middelburg[1,2].

1. Center for Clinical Transfusion Research, Sanquin Research; 2. Department of Clinical Epidemiology, Leiden University Medical Center; 3. Department of Haematology, HAGA teaching Hospital Den Haag.

ABSTRACT

Background: Haemato-oncology patients undergoing intensive chemotherapeutic treatments receive prophylactic platelet transfusions. Differences in count increments after transfusion of fresh or old platelets have been reported, but are difficult to translate directly into real clinical success of a transfusion. However, lower increments are used to label transfusions as "failed" and diagnose platelet transfusion refractoriness. Therefore, we now quantified the association of storage time with the expected percentage of failed transfusions, for a range of possible count increment thresholds, to estimate the number of unnecessary diagnoses of refractoriness.

Methods: Based on results from a meta-analysis, the expected percentages of failed and successful transfusions were estimated for two different definitions of fresh and old transfusions.

Results: For the 'Maximum storage 5 days' contrast (0-2 versus 3-5 days), based on the 24 hour absolute count increment, for thresholds ranging from 0 to 30, the difference in the percentages of failure, between old and fresh transfusions, ranged from 4.9% to 5.5%. Based on the 1 hour corrected count increment, for thresholds ranging from 0 to 15, the differences between old and fresh transfusions, ranged from 2.7% to 10.4%. After 24 hours these differences ranged from 4.3% to 6.2%.

Conclusion: Out of every 20 old platelet transfusions one will be considered failed, while a fresh platelet transfusion would have been successful. This will happen, irrespective of any patient characteristics or clinical factors. This failure is therefore likely to have limited clinical relevance and could result in an unnecessary diagnosis of refractoriness.

INTRODUCTION

Prophylactic platelet transfusions are an important supportive therapy for haemato-oncology patients undergoing intensive chemo-therapeutic treatments.[1] We recently performed a systematic review and meta-analyses, quantifying the association of platelet storage time and absolute and corrected count increments.[2] Our results confirmed the expected difference in count increments between fresh and old platelets. The observed difference in 1 hour corrected count increment was 2.11 (95% confidence interval (CI): 1.51 to 2.71) between fresh and old platelets. The difference in the 24 hour corrected count increment was 1.36 (CI: 0.12 to 2.60).[2] However, directly translating these differences into a clinically relevant interpretation is difficult. Especially since the relevance of platelet counts and count increments for the haemostatic effect, which is the true measure of success of a platelet transfusion, might be limited.[3-5]

One way in which a difference in count increment might become clinically relevant, completely independently of any potential effect on haemostasis, is by its influence on the diagnosis of refractoriness. What is mostly agreed upon is that a patient is to be considered refractory to platelet transfusion if he or she fails to show adequate increments in platelet count on at least two consecutive platelet transfusions.[1,6-11] Formally, these two consecutive transfusions are supposed to be both with fresh platelets (i.e. <72 hours of storage).[9,10] However, in clinical practice it is not possible to specifically order two consecutive transfusions of fresh platelets for all patients. Additionally, a blood bank supplying predominantly older platelets is likely to supply two consecutive old transfusions and a blood bank supplying predominantly fresh platelets is likely to supply two consecutive fresh transfusions. By basing the diagnosis of refractoriness on the perceived failure of two consecutive transfusions, while failure is defined based on count increments and count increments are known to depend on storage time, patients will be deemed refractory, while the storage time of the transfused product was really to blame. In these patients diagnostic work-up for suspected refractoriness will be started unnecessarily.

For the diagnosis of refractoriness the percentage of successful transfusions is more directly relevant than the observed absolute or corrected count increment. However, what constitutes a 'successful' or a 'failed' transfusion, based on count increments, is difficult to define exactly.[6,7,9] Different thresholds for what should be considered adequate count increments and corrected count increments have been suggested.[1,6,12] Some clinicians more informally consider a transfusion 'failed' if another one is needed the next day (i.e. no or clinically irrelevant 24 hour absolute count increment). Others calculate corrected count increments and strictly adhere to a certain pre-specified threshold for success of a transfusion. The exact definition chosen to determine the "success of a transfusion", based on platelet count derived measures, could affect the relative size of the effect of storage time on the percentage of successful transfusions and therefore on the number of unnecessary diagnoses of refractoriness.

Therefore, we now further investigated the previously reported count increments, to quantify the association of storage time with the expected percentage of failed transfusions, for a range of possible absolute and corrected count increment thresholds, to estimate the number of unnecessary diagnoses of refractoriness expected.

METHODS

We previously performed a systematic review and meta-analyses, including any publication indexed in MEDLINE (PubMed), EMBASE, Cochrane, CINAHL, Academic Search Premier, ScienceDirect and Web of Science databases, until February 2016, about the direct comparison of fresh versus old platelet transfusions and their effect on clinical measurements (i.e. platelet counts and derived measures) after transfusion. The terms 'fresh' and 'old' were analysed in different ways as described previously.[2] For the current analyses we selected two storage time contrasts to increase homogeneity in the definition of fresh and old platelets:

- **Maximum storage 5 days (0-2 versus 3-5 days):** Papers were included that reported results for zero to two days (fresh) and three to five days (old).
- **Extreme difference (0-2 versus 5-7 days):** Papers were included that reported results for zero to two days (fresh) and five to seven days (old).

The expected percentages of successful transfusions were estimated for the 1 hour and the 24 hour absolute and corrected count increments, for old and fresh transfusions. All reported absolute correct increment are expressed in [$\times 10^9$/l] and correct count increments are expressed in [/dm]. The percentages of success were derived from the normal distributions for these outcomes as estimated based on the weighted mean differences and standard deviations from the random effects models from the previously published meta-analyses[2] (for formulas see online supplemental material). Figure 1 shows the distributions of count increments for fresh and old transfusions. For each fixed threshold (x) the left area under the curve represents the percentage of failed transfusions and the right area represents the percentage of successful transfusions. The grey area represents the increase in the percentage of failed transfusions among transfusions of old platelets, compared to transfusions of fresh platelets.

Thresholds (x) for 'successful' or 'failed' transfusions varied from 0 to 30 for absolute count increments and from 0 to 15 for corrected count increments. Number need to tread (NNT) were calculated using the following formula:

$$NNT = \frac{1}{Absolute\ risk\ difference} =$$

$$= \frac{1}{P_{failure}(old) - P_{failure}(fresh)}$$

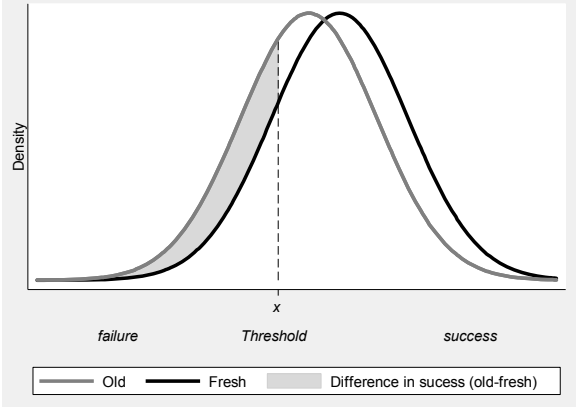


Figure 1: Distribution of platelet count increments

The area to the left of threshold represents failure and the area to the right of the threshold represents success. Grey area represents the difference between old and fresh distributions at the threshold x

RESULTS

Of the 46 papers selected in the original meta-analyses, 29 reported absolute or corrected count increments. The data of these 29 papers were used to estimate the distributions count increments and corrected count increments for fresh and old platelets and the percentages of 'successful' or 'failed' transfusions. Nine papers reported count increments: six papers reported 1 hour count increments of 4,822 transfusions, and eight reported 24 hour count increments of 3,531 transfusions. Twenty-seven papers reported corrected count increments: 23 reported the 1 hour corrected count increments of 19,117 platelet transfusions, and 23 reported the 24 hour corrected count increments of 8,032 platelet transfusions (Table 1).

Table 1 also shows the mean increment of old and fresh platelets and the combined standard deviations for each definition of old and fresh. Fresh platelets had higher mean absolute and corrected count increments than old platelets for all the contrasts studied.

Absolute Count increment

For the 'Maximum storage 5 days' contrast (0-2 versus 3-5 days), for thresholds ranging from 0 to 30 L, based on the 24 hour increment, the percentages of failed transfusions ranged from 32% to 66% for fresh platelets and from 37% to 71% for old platelets. This corresponded to differences, between old and fresh transfusions, ranging from 4.9% to 5.5% and NNT ranging from 18 to 20. Results for all thresholds are presented in table 2 and figure 2.

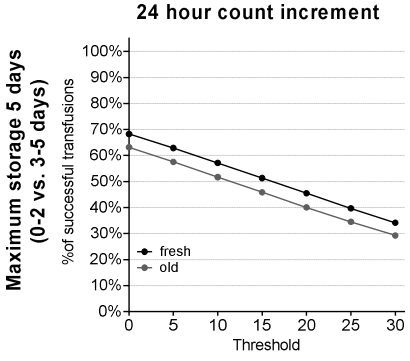


Figure 2: Percentage of successful transfusions as judged by the 24 hour absolute count increment, according to different thresholds for success

Table 1: Underlying distribution of fresh and old platelets and total number of studies and transfusions included in the analyses, according to different contrasts of old and fresh platelets.

Outcome Contrasts	Number of studies	Transfusions			Increment		
		Total	Fresh	Old	Mean fresh	Mean Old	Standard deviation*
24 hour absolute count increment							
Maximum storage 5 days (0-2 vs. 3-5 days)	5	3,063	581	2,482	16.15	11.47	33.97
1 hour corrected count increment							
Maximum storage 5 days (0-2 vs. 3-5 days)	15	18,049	4,113	13,936	14.32	12.21	8.03
Extreme difference (0-2 vs. 5-7 days)	10	6,693	3,341	3,352	13.93	11.24	6.54
24 hour corrected count increment							
Maximum storage 5 days (0-2 vs. 3-5 days)	15	6,813	2,165	4,648	8.26	6.91	8.76
Extreme difference (0-2 vs. 5-7 days)	8	2,393	1,003	1,390	8.78	7.43	7.29

*Combined for fresh and old

Corrected count increment

For the 'Maximum storage 5 days' contrast (0-2 versus 3-5 days), for corrected count increments thresholds ranging from 0 to 15, based on the 1 hour corrected count increment, the percentages of failed transfusions ranged from 4% to 54% for fresh platelets and from 6% to 64% for old platelets. This corresponded to differences, between old and fresh transfusions, ranging from 2.7% to 10% and NNT ranging from 37 to 10. After 24 hours these differences ranged from 4.3% to 6.2% and NNT ranged from 16 to 23. Results for all thresholds and for the 'Extreme difference' storage time contrast (0-2 versus 5-7 days) are presented in table 2 and figure 3.

DISCUSSION

As expected, we observed substantial differences in the percentage of failed and successful transfusions between fresh and old platelets. These results further indicate that between one out of 16 and one out of 37 transfusions with 3-5 day-old platelets will be considered failed while transfusions with 0-2 day-old platelets could have been successful. This two-and-a-half-fold difference is mostly due to the inclusion of results for the 1 hour corrected count increment, where the percentage of successful transfusions is influenced strongly be the chosen threshold. When considering 24 hour absolute or corrected count increments, numbers needed to treat were more stable between 16 and 23 (average 20), even for thresholds ranging from 0 to 30 for absolute count increments and from 0 to 15 for corrected count increments.

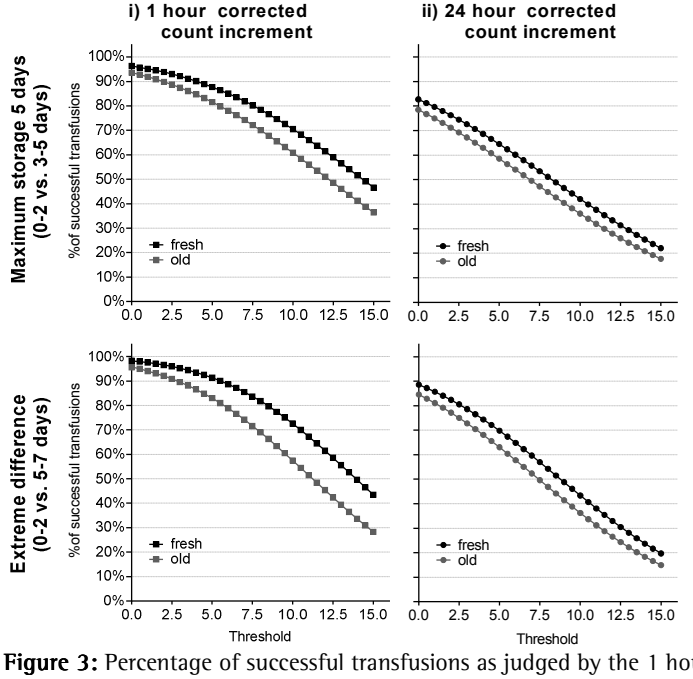


Figure 3: Percentage of successful transfusions as judged by the 1 hour and 24 hour correct count increment, according to different thresholds for success and different definitions of fresh and old platelets

Table 2: Percentage of failed transfusions judged by absolute and corrected count increments, according to different contrasts of fresh versus old platelets and different thresholds for success.

Absolute count increment				Corrected count increment			
Threshold	Fresh (% failed)	Old (% failed)	Difference (%)	Threshold	Fresh (% failed)	Old (% failed)	Difference (%)
1 hour increment, contrast: Maximum storage 5 days (0-2 vs. 3-5 days)							
0	NA	NA	NA	**0**	3.72	6.41	2.69
5	NA	NA	NA	**2.5**	7.04	11.32	4.27
10	NA	NA	NA	**5**	12.28	18.45	6.17
15	NA	NA	NA	**7.5**	19.77	27.86	8.09
20	NA	NA	NA	**10**	29.52	39.15	9.63
25	NA	NA	NA	**12.5**	41.03	51.44	10.41
30	NA	NA	NA	**15**	53.38	63.59	10.21
24 hour increment, contrast: Maximum storage 5 days (0-2 vs. 3-5 days)							
0	31.72	36.79	5.07	**0**	17.27	21.52	4.25
5	37.13	42.45	5.32	**2.5**	25.52	30.74	5.22
10	42.81	48.28	5.47	**5**	35.47	41.39	5.92
15	48.65	54.14	5.50	**7.5**	46.53	52.71	6.18
20	54.51	59.92	5.41	**10**	57.87	63.82	5.95
25	60.27	65.48	5.21	**12.5**	68.58	73.86	5.28
30	65.82	70.73	4.91	**15**	77.92	82.24	4.32
1 hour increment, contrast: Extreme difference (0-2 vs. 5-7 days)							
0	NA	NA	NA	**0**	1.66	4.28	2.62
5	NA	NA	NA	**2.5**	4.03	9.07	5.03
10	NA	NA	NA	**5**	8.62	16.99	8.37
15	NA	NA	NA	**7.5**	16.29	28.35	12.06
20	NA	NA	NA	**10**	27.41	42.46	15.05
25	NA	NA	NA	**12.5**	41.36	57.61	16.25
30	NA	NA	NA	**15**	56.51	71.70	15.20
24 hour increment, contrast: Extreme difference (0-2 vs. 5-7 days)							
0	NA	NA	NA	**0**	11.43	15.42	4.00
5	NA	NA	NA	**2.5**	19.45	24.96	5.51
10	NA	NA	NA	**5**	30.20	36.96	6.76
15	NA	NA	NA	**7.5**	43.02	50.39	7.38
20	NA	NA	NA	**10**	56.62	63.78	7.16
25	NA	NA	NA	**12.5**	69.48	75.66	6.18
30	NA	NA	NA	**15**	80.30	85.04	4.74

NA: not available, meta-analyses was not performed because less than 5 studies reported the outcome.

The true success of a platelet transfusion, should of course be measured by its haemostatic effect. It has been suggested that the relevance of storage time is very limited in this context.[13,14] This makes it even more worrisome that clinically relevant decisions, such as the decision to start diagnostic work-up for suspected platelet transfusion refractoriness, are still based on platelet count measurements, which do depend on storage time. If a blood bank supplies predominantly older platelets it would be likely to supply two consecutive old transfusions and a blood bank supplying predominantly fresh platelets would be likely to supply two consecutive fresh transfusions. Recipients from the 'old-supplier' are then likely to be deemed refractory one out of 16 to 23 times, where they would not have been considered refractory, if they had received transfusions from the 'fresh supplier'.

Being aware of this potential problem does not necessarily solve it. Clinicians might well be aware that two consecutively failed transfusions with older platelets do not necessarily indicate refractoriness to platelet transfusions. However, the mere fact that the two failed transfusions were with older platelets does not rule out refractoriness either. Therefore, out of precaution, every two consecutively failed transfusions should still be treated with similar

caution, even if the transfused platelets were 'old'. As a result, customers of a blood bank with predominantly older platelets are likely to start additional, unnecessary diagnostic work-up and raise unnecessary concerns in about one out of every twenty patients.

Similarly, seasonal differences in average storage time, or storage time differences related to blood groups could result in unnecessary concerns, since they are more likely to result in the transfusion of two consecutive old units. However, as mentioned above, knowing two units were old does not excuse a clinician from considering refractoriness for that patient. After all, the majority of transfusion failures occurring after transfusion of old units are completely storage time independent and therefore indicative of real refractoriness of the patient. After a single failed transfusion a clinician might still consider ordering a fresh unit for the next transfusion, if local blood supply logistics allow. However, if a patient is really refractory, any delay in diagnostic work-up will also result in a delay of appropriate treatment. Therefore, this option would be less preferable after multiple failed transfusions.

One solution for this problem could be to calculate increments corrected for storage time. However, for this measure to be useful in clinical practice would also require good consensus about which threshold should then be used to judge a transfusion as successful or failed.

Currently, there is not even consensus about the threshold for absolute and conventional corrected count increments. Reaching a consensus for the threshold for the "storage time-corrected count increment" would first mean reaching international consensus about the calculation of this measure. We therefore call experts in the field to suggest relevant calculations, simple enough to be applicable to daily clinical practice. In the meantime, the relatively large variation observed in estimates derived from 1 hour counts, might suggest the use of 24 hour increments, either absolute or corrected, to be preferable.

REFERENCES

1. Lee C, Ayob Y. Approach to managing platelet refractory patients. *ISBT Science Series.* 2015;10(S1):89-94.
2. Caram-Deelder C, Kreuger AL, Jacobse J, van der Bom JG, Middelburg RA. Effect of platelet storage time on platelet measurements: a systematic review and meta-analyses. *Vox sanguinis.* 2016;111(4):374-382.
3. Slichter SJ, Kaufman RM, Assmann SF, et al. Dose of prophylactic platelet transfusions and prevention of hemorrhage. *The New England journal of medicine.* 2010;362(7):600-613.
4. Stanworth SJ, Estcourt LJ, Powter G, et al. A no-prophylaxis platelet-transfusion strategy for hematologic cancers. *The New England journal of medicine.* 2013;368(19):1771-1780.
5. Kumar A, Mhaskar R, Grossman BJ, et al. Platelet transfusion: a systematic review of the clinical evidence. *Transfusion.* 2015;55(5):1116-1127; quiz 1115.
6. Bishop JF, Matthews JP, Yuen K, McGrath K, Wolf MM, Szer J. The definition of refractoriness to platelet transfusions. *Transfusion medicine (Oxford, England).* 1992;2(1):35-41.
7. Cook RJ, Heddle NM. Clinical trials evaluating pathogen-reduced platelet products: methodologic issues and recommendations. *Transfusion.* 2013;53(8):1843-1855.
8. Friedberg RC, Donnelly SF, Boyd JC, Gray LS, Mintz PD. Clinical and blood bank factors in the management of platelet refractoriness and alloimmunization. *Blood.* 1993;81(12):3428-3434.
9. Hod E, Schwartz J. Platelet transfusion refractoriness. *British journal of haematology.* 2008;142(3):348-360.
10. Schiffer CA. Diagnosis and management of refractoriness to platelet transfusion. *Blood reviews.* 2001;15(4):175-180.
11. Slichter SJ, Davis K, Enright H, et al. Factors affecting posttransfusion platelet increments, platelet refractoriness, and platelet transfusion intervals in thrombocytopenic patients. *Blood.* 2005;105(10):4106-4114.
12. Schiffer CA, Anderson KC, Bennett CL, et al. Platelet transfusion for patients with cancer: clinical practice guidelines of the American Society of Clinical Oncology. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2001;19(5):1519-1538.
13. Kreuger AL, Caram-Deelder C, Jacobse J, Kerkhoffs JL, van der Bom JG, Middelburg RA. Effect of storage time of platelet products on clinical outcomes after transfusion: a systematic review and meta-analyses. *Vox sanguinis.* 2017;112(4):291-300.
14. Triulzi DJ, Assmann SF, Strauss RG, et al. The impact of platelet transfusion characteristics on posttransfusion platelet increments and clinical bleeding in patients with hypoproliferative thrombocytopenia. *Blood.* 2012;119(23):5553-5562.

Appendix – percentage of successful transfusions: formulas

Given: i: the studies indicator

j: the fresh/old indicator

$\bar{x}_{ij}$: the mean in the j^{th} category (fresh/old) of the i^{th} study

SD_{ij}: the standard deviation in the j^{th} category (fresh/old) of the i^{th} study

n_{ij}: the sample size in the j^{th} category (fresh/old) of the i^{th} study

$SE_{ij} = SD_{ij} \times \sqrt{n_{ij}}$ by definition the standard error (SE) is the standard deviation (SD) times the square root of the sample size (n)

τ^2: the inter-study variation from the DerSimonian and Laird random effect model

From the individual studies we estimated the meta-analyses pooled mean ($\bar{x}_j$) and the standard deviation (SD_j) for the fresh and old platelets separately based on the random effects model, following the steps:

1. The estimate of the combined effect for heterogeneity is defined as the inverse of the variance: $w_j = \left(1/\sum({SE_{ij}}^2 + \tau^2)\right)$ (i.e. the weight of each study under the random effects model)
2. $SE_j = 1/\sum w_j$ by definition the SE is the inverse of the sum of the studies weights
3. $\bar{x}_j = (\sum \bar{x}_{ij} \times w_{ij})/\sum w_{ij}$ (i.e. the pooled effect size of each group)
4. $n_j = (\sum n_{ij} \times w_{ij})/\sum w_{ij}$ (i.e. the pooled sample size of each group)
5. $SD_j = SE_j \times \sqrt{n_j}$ (by definition)

The probability of success is given by: $P(X \geq threshold)$ where $X \sim N(\bar{x}_j, SD_j)$ and the pooled (fresh/old) standard deviation is: $SD = \Sigma(n_j - 1)s_j^2/\Sigma(n_j - 1)^{1/2}$

Reference

Bradburn MJ, Deeks JJ, Altman DG. 1998. sbe24: metan – an alternative meta-analysis command. command. *Stata Technical Bulletin* STB-44, pp.4-15.
Available at: http://www.stata-press.com/journals/stbcontents/stb44.pdf

CHAPTER 7

Continuing use of the terms prospective and retrospective and quality of reporting of observational studies: time to update the STROBE guideline?

International Journal of Epidemiology. 2016; 45 (2): 587-589.

Camila Caram-Deelder[1,2]; Aukje L. Kreuger[1,2];
Frits R. Rosendaal[2]; Johanna G. van der Bom[1,2];
Rutger A. Middelburg[1,2].

1. Center for Clinical Transfusion Research, Sanquin Research; 2. Department of Clinical Epidemiology, Leiden University Medical Center.

Clear and good reporting of observational studies[1-3] is essential to translate study's findings to daily practice and to allow correct inclusion of study results into synthesis of evidence, like meta-analyses.[4] The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines, published in 2007,[4-6] addressed an important aspect: a recommendation to refrain from using the terms prospective and retrospective or to clearly define what is meant by them. The use of these terms has caused much discussion in the past about the meaning of the words (for common definitions see supplemental material - appendix 1), the weight that is given to these terms and, as a consequence, their influence in the decision to fund or publish research.[4,7-9]

Six years after the STROBE publication the terms 'prospective' and 'retrospective' are still frequently used. In fact, in our perception it is difficult to find a clinical journal where these words are not every in the table of contents. This seems to suggest poor uptake of the STROBE recommendations and could therefore signal persistent poor quality of reporting. Therefore, in order to quantify to what extent authors follow the recommendation to refrain from using these terms and whether the use of these terms was associated with the overall quality of reporting (quantified as STROBE adherence score), we systematically reviewed 150 reports of observational studies in top clinical journals (general medicine, clinical specialist and general epidemiology journals). We also checked the frequency of use of the terms in PubMed. (for detailed methods see supplemental material - appendix 2)

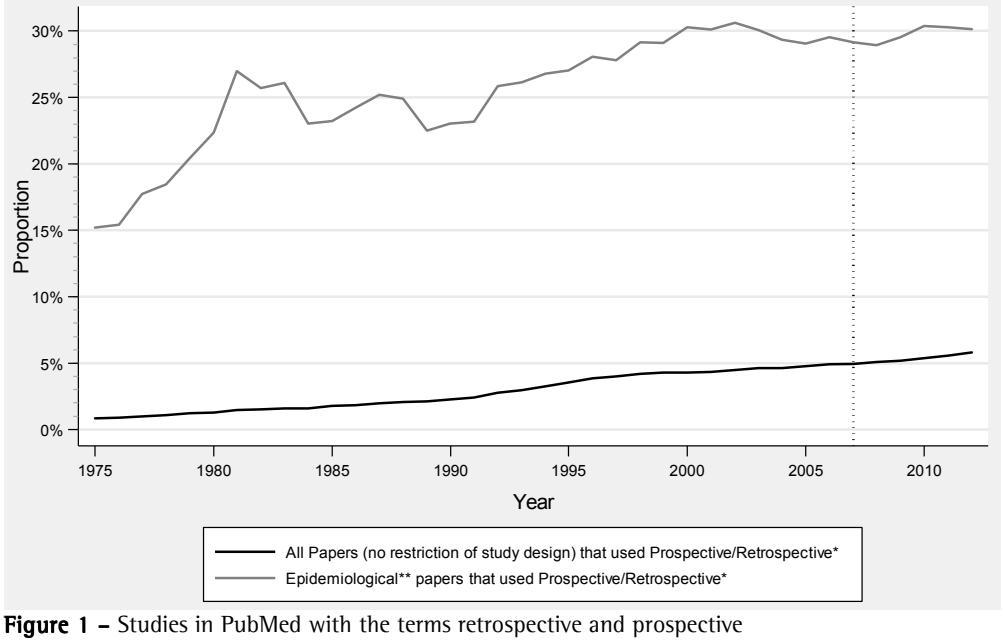


Figure 1 – Studies in PubMed with the terms retrospective and prospective
* Abstract or Title
** Epidemiologic Studies[Mesh] NOT Seroepidemiologic Studies[Mesh]
Searches restricted to papers with abstract available
The vertical dotted line corresponds to 2007, year of the STROBE guidelines publication

The terms 'prospective' and 'retrospective' were used in the title or abstract in 572,246 (4% of 13,862,297) papers indexed in PubMed from 1975 till 2013. Seventy-three percent of those papers (415,535) were indexed as epidemiologic studies (i.e. case-control, cohort, cross-sectional studies). The percentage of papers that use the terms 'prospective' and 'retrospective' increased over time; in 1975 one percent (831/ 100,458) of all papers and 15% (462/3,037) of the epidemiologic papers used these terms in the title or abstract; by 1995 these percentages increased to 4% (11,449/ 322,042) and 27% (9,114/ 33,704); and in 2013 six percent (52,575/851,631) of all papers and 31% (32,189/102,745) of the epidemiologic papers used the terms 'retrospective' and 'prospective' in the title or abstract. The latest percentage remained stable around 30% before and after the STROBE publication (figure 1).

In the leading clinical journals, in spite of the recommendation to refrain from using the terms, over half the selected papers (86/150) still uses the terms prospective or retrospective in the title, abstract or methods section. Only four of these papers define what they mean by them. Seventy four (49%) papers used the terms in the methods or abstract but not in the title. Thirteen (9%) papers used the terms exclusively in the abstract and 31 (21%) exclusively in the methods section. The terms were never used exclusively in the title. Fifty-five of the papers (37%) used the term in the title or abstract, this proportion is 5% higher than the proportion observed in PubMed (95% CI: -3% to 13%). Seventy-three papers (49%) used the terms prospective or retrospective in the methods sections. Twelve papers (8%) used the terms in the title and 53 (35%) in the abstract (figure 2).

Prospective was used 2.2 (95% CI: 1.5 to 3.2) times more often for reports of all study designs than retrospective. Papers published in general epidemiology journals used the terms less than papers published in clinical specialist and general medicine journals. Also the terms were used more to describe cohort studies than case-control and cross-sectional studies.

The use of the terms 'retrospective' and 'prospective', however, is not associated with the overall quality of the report, measured as STROBE adherence score neither any of its domains (*Setting, Participants, Variables*, *Data sources/measurement* and *Bias)*. Again no differences in quality of the report were observed according to papers using the terms or not, study type, journal recommending STROBE, journal type, and impact factor. (for detailed results see supplemental material - appendix 3)

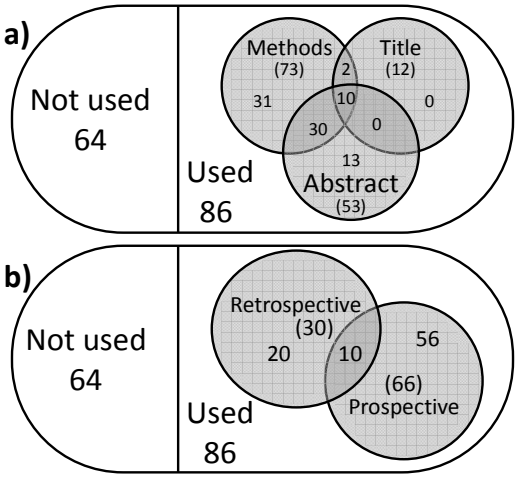


Figure 2 – Use of the terms prospective and retrospective
a) sections where the terms were used
b) terms used

In summary, we found that the use of the terms prospective and retrospective did not change after publication of the STROBE guideline suggesting no impact of STROBE on the use of these terms. The lack of any association of the use of these terms to the overall adherence to the rest of STROBE's advice further suggests this is not an intentional ignoring of STROBE's advice in general. Rather this seems specific to this particular part of STROBE's advice, possibly because authors are not aware of this advice.

We envision two possible actions. Either there should be an effort to make people aware of the advice to avoid the terms prospective and retrospective, or the STROBE should be updated and this advice should be removed. The decision which of these two options to follow should ultimately depend on the outcome of a discussion, on the merits and perils of using these terms, among a broad representation of the medical scientific community.

REFERENCES

1. Hoppe DJ, Schemitsch EH, Morshed S, Tornetta P, III, Bhandari M. Hierarchy of evidence: where observational studies fit in and why we need them. J Bone Joint Surg Am. 2009;91 Suppl 3:2-9.
2. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science - Supporting Information (Longer, more detailed version). PLoS Med. 2008;5(3):e67.
3. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. PLoS Med. 2008;5(3):e67.
4. Castillo RC, Scharfstein DO, Mackenzie EJ. Observational studies in the era of randomized trials: finding the balance. J Bone Joint Surg Am. 2012;94 Suppl 1:112-117.
5. Vandenbroucke JP, von Elm E, Altman DG et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology. 2007;18(6):805-835.
6. von EE, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-1457.
7. Dekkers OM, Egger M, Altman DG, Vandenbroucke JP. Distinguishing case series from cohort studies. Ann Intern Med. 2012;156(1 Pt 1):37-40.
8. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 3rd edition ed. Philadelphia: Wolters Kluwer; Lippincott Williams & Wilkins; 2008.
9. Vandenbroucke JP. Prospective or retrospective: what's in a name? BMJ. 1991;302(6771):249-250.

SUPPLEMENTAL MATERIAL

APPENDIX 1: Common definitions of the terms retrospective and prospective

Textbox – Common definitions of the terms retrospective and prospective

Early descriptions[1]
Prospective: Cohort
Retrospective: Case-control

Exposure measurement[2]
Prospective if the exposure measurement could not be influenced by the disease
Retrospective otherwise

Person-time[2]
Prospective: When the person-time accumulates after the study begins (exposure status is ordinarily recorded before disease occurrence)
Retrospective: When person-time accumulates before the study is conducted (even if the exposure status was recorded before the disease occurred)

Time the study begins[2]
Retrospective: historical events
Prospective: event concurrent with the study

1. Rothman KJ. Epidemiology: An Introduction. 3rd edition *New York, NY: Oxford University Press; 2012*Rothman KJ,
2. Greenland S, Lash TL. Modern epidemiology. 3rd edition ed. Philadelphia: Wolters Kluwer; Lippincott Williams & Wilkins; 2008.

APPENDIX 2: Methods

A systematic review of recently published observational papers in leading medical journals was performed. For these papers the use of the terms prospective and retrospective was recorded along with the quality of the report, measured as the adherence to the STROBE guideline. Also, for all papers indexed in PubMed, the use of the terms prospective and retrospective was quantified.

Selection of papers

Observational studies reporting original data from epidemiological studies in humans were included. The selection of papers was adapted from Knol et al.[1] This selection was based on a pragmatic sample of papers per chosen journal. It provides an overview of common practice in the most influential medical journals. In short: 150 papers, published in 20 journals divided into 3 groups:

(1) General medicine (GM; 5 journals, 10 papers each): Annals of Internal Medicine, British Medical Journal (BMJ), Journal of the American Medical Association (JAMA), Lancet, New England Journal of Medicine;

(2) General epidemiology (GE; 5 journals, 10 papers each): American Journal of Epidemiology, Epidemiology, International Journal of Epidemiology, Journal of Clinical Epidemiology, Journal of Epidemiology and Community Health; and

(3) Clinical specialist (CS; 10 journals, 5 papers each): American Journal of Respiratory and Critical Care Medicine, JAMA Psychiatry (formerly Archives of General Psychiatry), Arthritis and Rheumatism, Blood, Circulation, Clinical Infectious Diseases, Diabetes Care, Journal of the American Geriatrics Society, Journal of the National Cancer Institute, Paediatrics.

On the reference date (19 April 2013), the table of contents of the most recent issue of each journal was accessed. From these tables of contents, titles and abstracts were evaluated to select observational studies only (figure 1s). This procedure was repeated for previous issues, until the predefined number of papers per journal was selected. Finally, because all issues were reviewed entirely, more papers than necessary could be selected initially. Selected papers from the oldest issue were therefore included randomly up to the predefined number of papers.

The language of the publications was not taken into account because all chosen journals have exclusively English publications. Publications with no abstract were excluded. If eligibility could not be assessed based on the review of the title and abstract the full text was evaluated.

Data extraction

A data extraction form was developed to select papers (figure 1s) and to record the use of the terms prospective and retrospective and the quality of the report (measured as STROBE adherence score) (figure 2s). The STROBE adherence score was based on the methods section of the combined STROBE checklist for cohort, case-control, and cross-sectional studies.[2,3] Briefly: the five different domains – i) setting; ii) participants; iii) variables; iv) data sources/measurement; and v) bias – were analysed by the authors and structured into 31 individual items that covered each aspect addressed by the STROBE checklist. Next, an explanation of the items was given, based on the publication "STROBE explanation and elabo-ration document" (figure 2s). Items could be scored yes (1 point), partially (0.5 point), no (0 point), and not applicable. Papers' total score and domain scores were calculated as ten times the proportion of points scored from all applicable items (i.e. 10 x (0 x number of "no" + 0.5 x number of "partially" + 1 x number of "yes") / number of applicable items.

A remark about our motivations to develop our own tool: Different published tools to evaluate the quality of observational studies are available.[4-9] However, these are designed to evaluate the potential risk of bias. They therefore provide information about the quality of the study and not about the quality of the report.

Journal[1]:_____ **Vol/Issue:**_____

Ref ID[2]:		**First author:**		
1. Publication type: Original	☐ Not Original[3]	⇨ exclude	☐ Original	⇨ next question
	☐ Abstract not available	⇨ exclude	☐ Short report	⇨ next question
2. Observational (not experimental)	☐ No	⇨ exclude	☐ Yes	⇨ next question
3. Compare groups	☐ No	⇨ exclude	☐ Yes	⇨ next question
4. Type of study[4]	☐ Case-control	☐ Cross-sectional	☐ Cohort	☐ Other:
5. Human[5]	☐ No	⇨ exclude	☐ Yes	⇨ next question
6. Comment:				
7. To be included?	☐ **No**		☐ **Yes**	

Figure 1s – Inclusion form

If there is any doubt the paper should be include for the next step (full text), it can be exclude then.
1. code from journal list;
2. all information about the paper (title, journal name, issue, etc) is storage in the database; This number does not change after the selection (same number from selection to final analyses);
3. Letter, review, case/series report, table of contents, educational, pictures, poems, comments etc.
4. Reference: STROBE definitions
5. Exposure/outcome/unit of measurement

We were interested in the quality of the report rather than in quality of the studies. As recognized by the STROBE initiative, there are two distinct requirements to allow a reader to judge the quality of a study. First, the reader will need subject matter knowledge. Second, the report should be of sufficient quality. The first requirement is not related to or under the influence of the study or the authors of the report and was not the topic of the present study. We focussed entirely on the quality of the report. The consensus of what is necessary for a high quality report, was published and minutely explained by the STROBE initiative[2,3]. Therefore, we chose to use the STROBE recommendation to quantify the quality of reporting even though this, in itself, has absolutely no bearing on the quality of the study.

Additionally it was verified if the STROBE guideline was mentioned in any section of the paper. If the methods referred to a previous paper, the information from that paper was also used. Papers' selection and data extraction were performed by two reviewers independently. Disagreements were discussed with a third reviewer to reach a consensus.

For each journal the journal type (general medicine, general epidemiology and clinical specialist), impact factor, and STROBE recommendation were also recorded. Journals' impact factors (2012) were obtained from the journals' website. STROBE recommendation was defined as 'yes' if the STROBE guideline was recommended by the journal in the "recommendations to authors" in its website and 'no' otherwise.

Use of the terms in journals indexed in PubMed

A PubMed search was performed to create an overview of changes in the use of the terms prospective and retrospective over time. The total number of papers indexed in PubMed and the number of papers that used the terms in the title and abstract were both quantified per year from 1975 till 2013. Epidemiologic studies were selected using Medical Subject Headings (MESH) terms, excluding *"seroepidemiologic studies"*, included as epidemiologic study in the MESH term. The proportion of papers using the terms was calculated both for all papers and for the subgroup of e*pidemiologic studies*. Only papers with an abstract available were selected in the PubMed searches.

Statistical

All statistical analyses were carried out in STATA version 12. Results are reported as proportions and differences between proportions (with 95% confidence intervals) for use of the terms prospective and retrospective and median and interquartile range for STROBE adherence scores (total and domain scores).

List of papers

The complete list of 150 papers included in the analyses can be found at: https:// goo.gl/iHdnWt or https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyv335 (Supplementary data)

Journal:_____ Vol./Issue:_____ First Author:_____ Ref id:_____

Inclusion/Exclusion of the paper

01. After read full text/methods: ☐ Include ☐ Exclude; Reason (inclusion form):

Study design

02. According to **authors**: ☐ Not defined ☐ Case-control ☐ Cohort ☐ Cross-sectional ☐ Other:
03. According to **reviewers**: ☐ Not defined ☐ Case-control ☐ Cohort ☐ Cross-sectional ☐ Other:

Prospective/Retrospective

04. Are the terms used to describe the paper? [0] No [1] Prospective [2] Retrospective [3] Both
05. If yes, are the terms defined? [0] No [1] Yes [9] Not applicable
06. If yes, definition (highlight in the paper): [0] Not given [9] Not applicable

FROM the methods section of the combined STROBE checklist for cohort, case-control, and cross-sectional studies
[0] No [1] Partially [2] Yes [9] Not applicable

Domain setting

Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection.

07. Setting - Type of recruitment or source of selection site (e.g., outpatient clinic, cancer registry, or university hospital).	0	1	2	9
08. Locations - Refer to the countries, towns, hospitals or practices.	0	1	2	9
09. Periods of recruitment - Date (or date range) when subjects were recruited or invited to participate of the study.	0	1	2	9
10. Periods of exposure - Date (or date range) when exposure occurred, relative to recruitment and outcome.	0	1	2	9
11. Periods of follow-up - Date (or date range) of follow-up, relative to recruitment and outcome.	0	1	2	9
12. Periods of data collection - Date (or date range) when data was collected or measurements made for the study.	0	1	2	9

Domain participants

(a) Cohort: Give the eligibility criteria, and the sources and methods of selection of participants.
Describe methods of follow-up.

13. Eligibility criteria of participants - Eligibility criteria may be presented as inclusion and exclusion criteria.	0	1	2	9
14. Sources of selection of participants - What is the population that participants come from.	0	1	2	9
15. Methods of selection of participants - How (method) participants were selected in the population (e.g. referral or self-selection through advertisements, all patients in the hospital database, random sample). Also response rate if applicable.	0	1	2	9
16. Methods of follow-up - How the follow up was measured (e.g. mail questioner, home visit, hospital visit, mortality database).	0	1	2	9

(a) Case-control: Give the eligibility criteria, and the sources and methods of case ascertainment and control selection.
Give the rationale for the choice of cases and controls.

17. Eligibility criteria of cases - Eligibility criteria may be presented as inclusion and exclusion criteria.	0	1	2	9
18. Sources of case ascertainment - What is the population where cases come from.	0	1	2	9
19. Methods of case ascertainment - How (method) the cases were selected in the population (e.g. referral or self-selection through advertisements, all patients in the hospital database, random sample). Also response rate if applicable.	0	1	2	9
20. Control selection sources - What is the population where controls come from (e.g. registry, general population, outcome free patients in a hospital)	0	1	2	9
21. Control selection methods - How (method) the controls were selected in the population (e.g. referral or self-selection through advertisements, all patients in the hospital database, random sample). Also response rate if applicable.	0	1	2	9
22. Give the rationale for the choice of cases and controls - Explanation why those controls were chosen for the study.	0	1	2	9

(a) Cross-sectional: Give the eligibility criteria, and the sources and methods of selection of participants.

23. Eligibility criteria - Eligibility criteria may be presented as inclusion and exclusion criteria.	0	1	2	9
24. Sources of selection of participants - What is the population that participants come from.	0	1	2	9
25. Methods of selection of participants - How (method) participants were selected in the population (e.g. referral or self-selection through advertisements, all patients in the hospital database, random sample). Also response rate if applicable.	0	1	2	9

(b) IF MATCHED (also for cross-sectional if matched)
Cohort: For matched studies, **give** matching criteria and **number of exposed and unexposed**
Case-control: For matched studies, **give** matching criteria and the **number of controls per case**

26. Matching criteria - Give a list of variables used for the matching criteria.	0	1	2	9
27. Number of "exposed and unexposed" or "controls per case" - Rate.	0	1	2	9

Domain variables

Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers.
Give diagnostic criteria, if applicable

28. Outcomes	Authors should clearly state outcomes as outcomes, confounders as confounders etc. It is common papers give a list of variables used in the study but don't make any differentiation between witch one is the outcome, exposure, potential confounders or effect modifiers. Later variables are used (for example) to "adjusted for" in multivariate models, but no statement was made that those variables were potential confounders.	0	1	2	9
29. Exposures or predictors		0	1	2	9
30. Potential confounders		0	1	2	9
31. Potential effect modifiers		0	1	2	9
32. Diagnostic criteria	Disease outcomes require detailed description of the diagnostic criteria. This applies to criteria for cases in a case-control study, disease events during follow-up in a cohort study and prevalent disease in a cross-sectional study. (e.g. anaemia should be follow by the haemoglobin level or clinical evaluation of the (listed) signs/ symptoms.	0	1	2	9

Domain data sources/measurement

For each variable of interest, **give** sources of data and details of methods of assessment (measurement).
Describe comparability of assessment methods if there is more than one group.

33. Sources of data - Where de data come from (e.g. patient charts, files, questionnaire).	0	1	2	9
34. Details of methods of assessment (measurement) - How the data was measured, what methods were used, including details of the reference standard that was used; (e.g. laboratory test details).	0	1	2	9
35. Comparability of assessment methods - If sources and methods are not the same the comparability (full or partial) should be made clear, full information should be given for readers judge how good (or bad) it is.	0	1	2	9

Domain Bias

Describe any efforts to address potential sources of bias.

36. Measurement/Information bias - Describe any effort to measure (get information from) cases and controls in the same way.	0	1	2	9
37. Selection bias - Describe any effort to control (or check) if the probability of including cases or controls was associated with exposure.	0	1	2	9

Figure 2s – Data extraction form – comments definitions and explanations

APPENDIX 3: Additional results

Selection of papers

From the 1,125 items listed in the table of contents of 67 issues of the 20 journals, 356 were excluded because they did not have an abstract available. Five hundred and four of the remaining 769 (66%) were excluded based on the titles and abstracts. Of the remaining 256 initially selected papers, the full text of 172 was assessed until the predefined sample size of 150 papers was reached – i.e. 22 papers initially considered eligible were excluded upon review of the full text (figure 3s).

STROBE recommendation

Eight out of 20 journals (40%) recommended the use of the STROBE guideline: 2/5 general epidemiology, 3/5 general medicine, and 3/10 clinical specialist. The median impact factor of journals that did recommend STROBE was 13.0 (IQR: 6.4 to 16.0), and for journals that did not recommend STROBE it was 8.4 (IQR: 5.4 to 14.0), Kruskal-Wallis rank test p-value 0.589.

STROBE was mentioned by the authors in one paper. This paper was published in a journal that does not recommend the use of the STROBE guideline.

Use of the terms prospective and retrospective in leading clinical journals

Table 1s shows the proportion of papers that used the terms prospective and retrospective, according to study type, STROBE recommendation, journal type, and impact factor. The terms prospective and retrospective were used 31% more often (95% CI: 14% to 47%) to describe cohort studies (68/101) than to describe case-control (11/30) and cross-sectional studies (7/19). Among the 65 cohort studies that used the terms 75% used only the term prospective, 20% used only the term retrospective, and 5% used both terms. Of the six case-control studies that used the terms, 50% used only the term prospective, 17% used only the term retrospective, and 33% used both terms. Of the six cross-sectional studies using the terms 33% used only prospective, 33% used only retrospective, and 33% used both.

Of the papers published in journals that recommended using the STROBE 58% (49/85) used the terms. In journals which do not recommend STROBE this was 57% (37/65). The terms prospective and retrospective were used in the title, abstract or methods section in 48% (24/50) of papers published in general epidemiology journals, and 62% (31/50) of papers published either in clinical specialists or in general medicine journals. Papers published in general epidemiology journals used the terms 14% less often (95% CI: -31% to 3%) than papers published in clinical specialist and general medicine journals (table 1s).

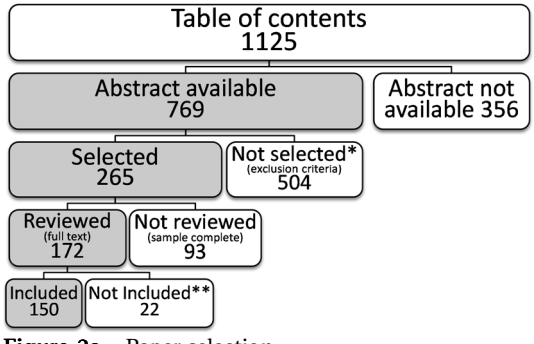


Figure 3s – Paper selection
*"Not original" or "Not Observational" or "Not compare groups" or "Not in humans"
**met one or more exclusion criteria after text full access

In journals in the highest quartile of impact factor the terms were used in 60% (24/42) of the papers, in the second quartile 40% (14/35), in the third quartile 67% (20/30), and in the fourth quartile 62% (28/45) used the terms (table 1s).

Among the 73 papers that used the terms in the methods section to describe the study design 4 (5%) provided a definition of the terms. Three defined the terms as a "*future/past calendar time frame*" and one defined the term "retrospective" as "*historical-cohort*".

STROBE adherence score

For the 150 papers reviewed the inter-observer consistency of the STROBE adherence score, calculated as interclass correlation (ICC) for ordinal variables, was 0.76 (95%CI 0.74 to 0.78) classified as "good". The inter-reviewer variation, also calculated as ICC for continuous variables, for score total was 0.81 (95%CI 0.73 to 0.86), and for domains it was 0.87 (95%CI 0.84 to 0.88), both classified as "good".

Table 1s shows the STROBE adherence score for papers using the terms prospective and retrospective, for papers not using these terms, and for all papers together according to study type, STROBE recommendation, journal type, and impact factor. The median of the STROBE adherence score was 8.1 (IQR 7.5 to 8.7). Papers that used the terms prospective and retrospective had a median score of 8.0 (IQR 7.5 to 8.6). Papers that did not use the terms had a median score of 8.3 (IQR 7.4 to 8.9). Cohort studies had a median score of 8.1 (7.6 to 8.7), case-control studies of 7.9 (IQR 7.2 to 8.6) and cross-sectional studies of 8.0 (IQR 7.2 to 8.9). In journals that recommended STROBE papers had a median score of 7.9 (IQR 7.4 to 8.6) and in journals that do not recommend STROBE the median was 8.1 (IQR 7.5 to 8.8). Among papers published in general epidemiology journals the median was 8.6 (IQR 8.1 to 8.9); for clinical specialist journals 7.8 (IQR 7.2 to 8.4) and for general medicine journals 7.9 (IQR 7.5 to 8.4). in journals in the highest quartile of the impact factor the median score was 8.6 (IQR 7.8 to 8.8), in the second quartile 8.5 (IQR 7.7 to 8.9), in the third quartile 7.8 (IQR 7.2 to 8.3), and in the fourth quartile 7.9 (IQR 7.4 to 8.3). Results were similar whether the terms prospective and retrospective were used or not (table 1s).

Similar results were observed for each one of the domains of the STROBE adherence score. The median score for the domain Setting was 9.2 (IQR 8.3 to 10.0); for the domain Participants 10.0 (IQR 9.4 to 10.0); Variables 7.0 (IQR 6.0 to 8.0); Data sources/measurement 10.0 (IQR 10.0 to 10.0) and Bias 2.5 (IQR 0.0 to 5.0). Results for the domain scores were similar according to study type, STROBE recommendation, journal type, and impact factor, and whether the terms prospective and retrospective were used or not (figure 4s and table 2s).

Papers that did not use the terms prospective and retrospective had higher median STROBE adherence scores only in the subgroup of papers in general medicine journals that recommend STROBE and general epidemiology journals (independent of the STROBE recommendation, figure 5s).

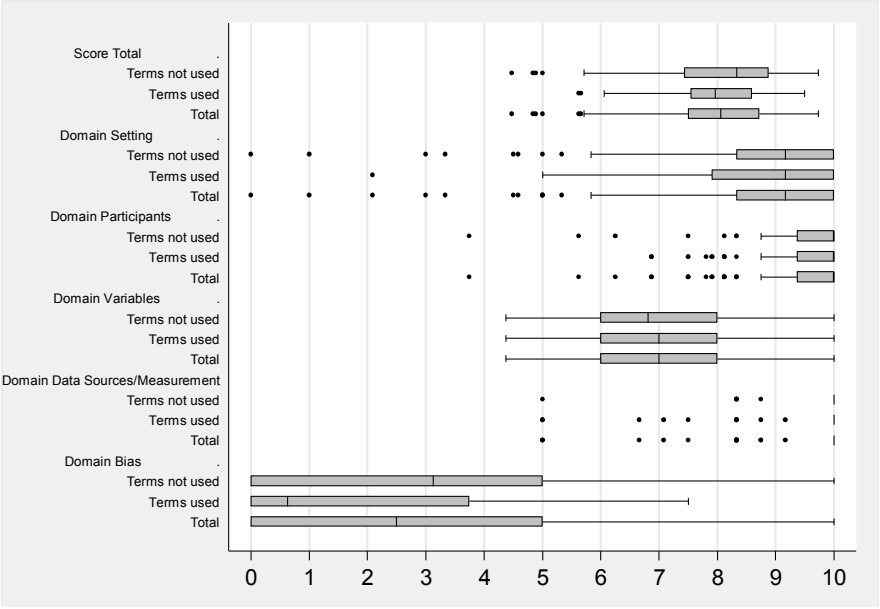


Figure 4s – STROBE adherence score, total and for separate domains according to use of the terms prospective and retrospective

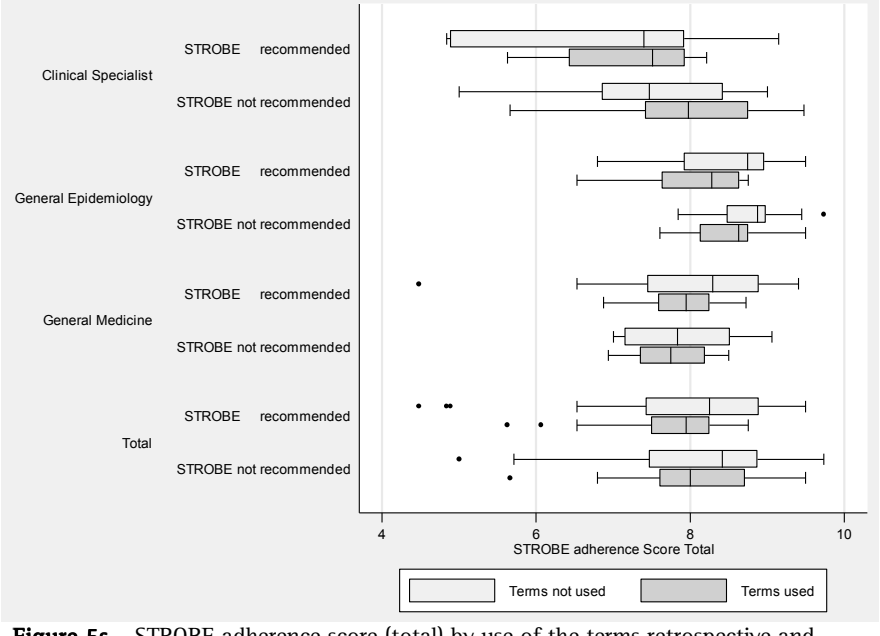


Figure 5s – STROBE adherence score (total) by use of the terms retrospective and prospective in the paper title, abstract or method section; strobe recommendation and journal type

Table 1s – Use of the terms retrospective and prospective

	Terms used[a]			Terms not used[a]			Difference proportion terms used (95%CI)	STROBE adherence Score Total, median (IQR[b])
	n	%	STROBE adherence Score Total, median (IQR[b])	n	%	STROBE adherence Score Total, median (IQR[b])		
Total	86	57%	8.0 (7.5 to 8.6)	64	43%	8.3 (7.4 to 8.9)	57% (49% to 65%)[c]	8.1 (7.5 to 8.7)
Study Type								
Cohort	68	67%	8.0 (7.6 to 8.6)	33	33%	8.5 (7.5 to 8.9)	reference	8.1 (7.6 to 8.7)
Case-control	11	37%	7.6 (7.2 to 8.1)	19	63%	8.1 (7.0 to 9.0)	31% (11% to 50%)[d]	7.9 (7.2 to 8.6)
Cross-sectional	7	37%	8.0 (7.2 to 8.7)	12	63%	8.3 (7.3 to 8.9)	30% (7% to 54%)[d]	8.0 (7.2 to 8.9)
STROBE recommendation								
Not Recommended	37	57%	8.0 (7.6 to 8.7)	28	43%	8.4 (7.5 to 8.9)	reference	8.1 (7.5 to 8.8)
Recommended	49	58%	7.9 (7.5 to 8.3)	36	42%	8.2 (7.4 to 8.9)	-1% (-17% to 15%)[d]	7.9 (7.4 to 8.6)
Journal type								
General Epidemiology	24	48%	8.6 (7.9 to 8.7)	26	52%	8.8 (8.5 to 9.0)	reference	8.6 (8.1 to 8.9)
Clinical Specialist	31	62%	7.9 (7.4 to 8.4)	19	38%	7.4 (6.5 to 8.4)	-14% (-33% to 5%)[d]	7.8 (7.2 to 8.4)
General Medicine	31	62%	7.9 (7.5 to 8.3)	19	38%	8.2 (7.3 to 8.9)	-14% (-33% to 5%)[d]	7.9 (7.5 to 8.4)
Impact Factor (2012)								
1st quartile	24	60%	8.3 (7.8 to 8.7)	16	40%	8.7 (7.8 to 9.0)	reference	8.6 (7.8 to 8.8)
2nd quartile	14	40%	8.6 (7.8 to 8.7)	21	60%	8.5 (7.6 to 8.9)	20% (-2% to 42%)[d]	8.5 (7.7 to 8.9)
3rd quartile	20	67%	7.8 (7.4 to 8.3)	10	33%	7.6 (6.5 to 8.5)	-7% (-29% to 16%)[d]	7.8 (7.2 to 8.3)
4th quartile	28	62%	7.9 (7.4 to 8.1)	17	38%	8.2 (7.4 to 8.8)	-2% (-23% to 19%)[d]	7.9 (7.4 to 8.3)

a. retrospective and prospective in the paper title, abstract or method section; b. interquartile range (Q1 to Q3); c H_0: prop(1) = 0.5; d H_0: prop(1=reference)-prop(2) = 0

Table 2s – STROBE adherence score, total and for separate domains

	n	Median (IQR)*					
		Score Total	Setting	Participants	Variables	Data sources/ measurement	Bias median
Total	150	8.1 (7.5 to 8.7)	9.2 (8.3 to 10.0)	10.0 (9.4 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	2.5 (0.0 to 5.0)
Use of the terms**							
Terms not used	64	8.3 (7.4 to 8.9)	9.2 (8.3 to 10.0)	10.0 (9.4 to 10.0)	6.8 (6.0 to 8.0)	10.0 (10.0 to 10.0)	3.1 (0.0 to 5.0)
Terms used	86	8.0 (7.5 to 8.6)	9.2 (7.9 to 10.0)	10.0 (9.4 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	0.6 (0.0 to 3.8)
Study Type							
Case-control	30	7.9 (7.2 to 8.6)	9.1 (7.9 to 9.2)	8.8 (8.1 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	2.5 (0.0 to 5.0)
Cohort	101	8.1 (7.6 to 8.7)	9.2 (8.3 to 10.0)	10.0 (10.0 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	2.5 (0.0 to 5.0)
Cross-sectional	19	8.0 (7.2 to 8.9)	9.5 (9.0 to 10.0)	10.0 (10.0 to 10.0)	6.0 (5.5 to 8.0)	10.0 (10.0 to 10.0)	2.5 (0.0 to 5.0)
STROBE recommendation							
Not Recommended	65	7.9 (7.4 to 8.6)	9.2 (8.3 to 10.0)	10.0 (8.8 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	1.3 (0.0 to 5.0)
Recommended	85	8.1 (7.5 to 8.8)	9.2 (8.3 to 10.0)	10.0 (9.4 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	2.5 (0.0 to 5.0)
Journal type							
Clinical Specialist	50	7.8 (7.2 to 8.4)	8.8 (6.7 to 10.0)	10.0 (8.8 to 10.0)	6.3 (6.0 to 8.0)	10.0 (10.0 to 10.0)	2.5 (0.0 to 3.8)
General Epidemiology	50	8.6 (8.1 to 8.9)	9.3 (9.0 to 10.0)	10.0(10.0 to 10.0)	7.3 (6.5 to 8.5)	10.0 (10.0 to 10.0)	3.8 (1.3 to 5.0)
General Medicine	50	7.9 (7.5 to 8.4)	9.1 (7.9 to 9.2)	10.0 (8.8 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	0.6 (0.0 to 5.0)
Impact Factor (2012)							
1st quartile	40	8.6 (7.8 to 8.8)	9.3 (9.2 to 10.0)	10.0 (10.0 to 10.0)	7.0 (5.9 to 8.3)	10.0 (10.0 to 10.0)	2.5 (0.0 to 5.0)
2nd quartile	35	8.5 (7.7 to 8.9)	9.2 (8.3 to 10.0)	10.0 (9.4 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	2.5 (0.0 to 5.0)
3rd quartile	30	7.8 (7.2 to 8.3)	8.8 (7.1 to 9.6)	10.0 (8.8 to 10.0)	6.3 (6.0 to 8.0)	10.0 (10.0 to 10.0)	1.9 (0.0 to 3.8)
4th quartile	45	7.9 (7.4 to 8.3)	9.2 (7.5 to 9.2)	10.0 (8.8 to 10.0)	7.0 (6.0 to 8.0)	10.0 (10.0 to 10.0)	0.0 (0.0 to 5.0)

* interquartile range (Q1 to Q3) ** prospective and retrospective in the paper title, abstract or method section

CHAPTER 8

Summary and general discussion

This thesis explores the potential of secondary data (i.e. routine clinical data and data for meta-analysis) to answer questions regarding safety and efficacy of blood products. That is, this thesis derives knowledge by using data that were generated for a different purpose. Despite of the limitations that secondary data carry, mostly related to data quality, it has the advantage of making it possible to study safety and efficacy of medical interventions also in large sample sizes.

The manner how secondary data are recorded can (and often does) differ from what researchers would have chosen to record if the data were generated having as their purpose a specific research question. Therefore, before using secondary data to answer any research question, all recorded data must be validated. Validation is crucial in any research using secondary data.

Validation is often laborious and involves extensive investigation into the meaning and source of the data. Validated secondary (*""retrospective""*) data, when considered of good quality, can be as good as (*""prospective""*) primary data. When reporting on secondary data researchers must clarify the validity and limitations of the data and results.[1]

All these topics are essential in both questions addressed in this thesis:

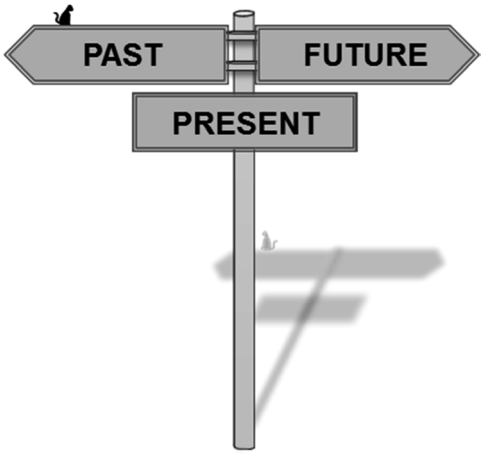
Should blood supplies take into account the sex of the red blood cell donors? (**chapter 2**)

How long is it safe and efficient to store platelet products? (**chapters 3 to 6**)

Are the terms 'retrospective' and 'prospective' necessary to describe observational studies? (**chapter 7**)

To answer these questions, data as a source of information and knowledge were used in several ways as described in the introduction (chapter 1). This final chapter presents recommendations on how to deal with validity of secondary data. It also gives a **summary** about the studied topics:

- **Past**: what was already known before this thesis
- **Present**: what this thesis add for each of the topics studied
- **Future**: principal implications and recommendations



VALIDITY

The path from data to knowledge is not straightforward. In the path *Data-Information-Knowledge Hierarchy*, before information can be translated into knowledge data needs to be transformed into information. When working with secondary data, either using data from a data warehouse or in a meta-analysis, one of the biggest challenges is merging. A deep understanding of the data is needed before merging. In the modelling phase there is another challenge, which is to consider, incorporate and validate the relationships between the variables that surround exposure and outcome. These two challenges, merging and modelling, will be discussed below.

Merging studies

Meta-analyses are techniques to merge and contrast results (i.e. data) from multiple studies. It consists of identifying patterns or sources of disagreement among the results of studies and when possible summarise as estimators. The main difference between systematic review and meta-analysis is the summarisation of results as estimators. In other words, systematic reviews answer research questions by collecting and summarising all empirical evidence that fits pre-specified eligibility criteria. A meta-analysis is the use of statistical methods to summarise the results of these studies.[1-5]

Validation consists of first judging to what extent populations and studies are comparable and can consequently be pooled (i.e. summarized as estimators). Secondly, validation is performed to ensure that the extracted data are ready to be pooled via statistical methods. Several pitfalls can occur during the validation process. Some of them are presented in the table 1, along with examples and solutions.

Besides the solutions presented in tables 1 researchers should always try to obtain extra information by contacting primary authors. Finally, a sensitivity analysis should always be performed to verify the influence of the decisions made on the final result.

Merging databases

Merging different source databases is a key step in studies where different sources are used. In our studies the sources were 6 (chapter 2) and 10 (chapter 7) different hospital databases that had their own (customised) computer platforms, codes and routines to collect and store information. A query was developed in one hospital and afterwards adapted to retrieve the information in other participating hospitals with the same structure.

The content of each database was individually described and translated to a common format to allow merging information from different hospitals. This task was labour intensive, delicate and in some occasions surprising. A simple example of this is the distinct codes that contain the same information across hospitals such as sex of the patient coded as

'M/F', 'M/V, 'Man/Vrow' or '0/1'.

Another, slightly more complex, example is the patient blood group (ABO and Rhesus D). This information was recorded in a single variable coded as, for example, 'A-/A+/B-/B+/O-/O+/AB-/AB+', or 'A neg/A pos/B neg/B pos/O neg/O pos/AB neg/AB pos'. Or in two variables that, when combined, contain the information about the blood group. For example: 'A/B/O/AB' and '-/+' or

'A/B/O/AB' and 'negative/positive'

Those examples are minor issues when merging datasets but still must be handled with attention and patience or the information can be lost along the way.

Finally, a highly complex example that could lead to major issues: unique product identification codes. As mentioned before, hospitals have their own system and the retrieved data revealed their own particularities. In the Netherlands the blood unit number is called, "Eenheid Identificatie Nummer", or "eenheidnummer" (EIN). EIN is a unique number assigned at the time of the donation. It is printed and fixed to each part of the satellite donation bag. This EIN is set by the national Dutch blood bank (Sanquin). Figure 1 shows the diagram of the EIN from donation to transfusion (or factory in the case of plasma). Red blood cells and products donated by apheresis carry the donation EIN until the final component. Pooled platelets (derived from whole blood donations) get a new EIN number during component production process. Plasma fraction of whole donations are sent to the factory under the donation EIN and lose this code during plasma product production.

Table 1 - Common pitfalls, examples and solutions of data validity in meta-analyses

Pitfall	Example	Our solution
Studies do not report variability	Some studies, specially old studies, do not report any measure of variability.	Consider imputation.
Results are showed only as graphs	Only graphs but no numbers reported as text in the manuscript.	Extract data using image reading scan (analogous to measure with a ruler but more accurate).
Heterogeneity between population	All but one study report results stratified (female vs. male). one study reports only total population.	Consider not to include the study.
Different estimations	Studies differ about the reported estimators: mean, median, odds ratios, risk ratio, absolute numbers, raw ratios, percentages, standard error, standard deviation, etc.	Recalculate: there are well established techniques to recalculate estimators.
Different scales	Study 1 reports: all platelets were transfused within 24 hours of donation. Study 2 reports: all platelets were transfused at day 1 of storage.	Recalculate. Verify the definition of "1" day of storage in relation to 24 hours of donation. In the Netherlands for example donation is considered day 0.
Low rates or no events	"no events (for example bleeding) were observed during follow up".	Consider to adjust estimations. There are well established techniques for adjustments.
Different time frame	Outcome: count increment after platelet transfusion Study 1 reports 8h count increment Study 2 reports 4h and 12h count increment	Consider interpolation.
Different outcome measurement	Study 1 reports number of blood units per 10 patient days Study 2 reports number of blood units per patient per treatment course	Consider recalculation only if the distribution of units over treatment course is known
Different definitions	Study 1: fresh: 2-5 vs old: 6-7 days Study 2: fresh: 1-4 vs old: 5-7 days Study 3: fresh: 1-2 vs old: 4-5 days Study 4: report results for each storage day from 1 to 7 (no dichotomization)	Present results for different definitions of fresh and old, for example: a) original definition (as reported): all studies included b) maximum storage 5 days (0-2 vs. 3-5): studies 3 and 4 included c) extreme difference (0-2 vs. 5-7): studies 3 and 4 included d) extended storage (0-5 vs. 6-7): studies 1 and 4 included

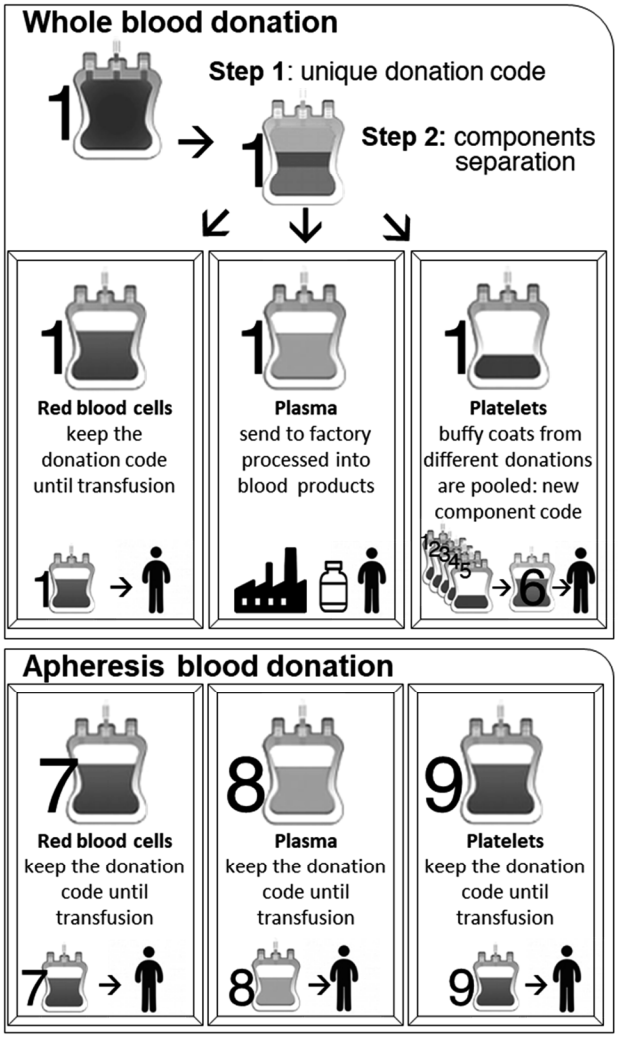


Figure 1 – diagram of the EIN from donation to transfusion or factory

Numbers 1 to 9 represent unique EIN codes.

Numbers 1 to 5 represent EIN given at whole blood donations.

Number 6 represent a new EIN given to the new component after pooling of buffy coats and medium.

Numbers 7 to 9 represent apheresis donation EIN.

Plasma fraction of whole donation is send to factory under the donation EIN and lose this code during products production.

EINS are formed by 13 or 16 characters: the letter N followed by 12 or 15 digits. The EIN can be broken down into the following blocks:

- Characters 1 to 5 (the letter N + 4 digits) represent the production centre (or supplier);
- Characters 6 and 7 represent the year in which the blood was donated;
- Characters 7 to 13 are the specific identification number, or internal number;
- Characters 14 to 16 bring information about production process. They are only used during the production period (from donation to final components) and discarded after production.

Figure 2 shows an example of a blood bag label. The label brings among others, the information of supplier, donation method, day and time, expiring date and time, ABO Rh D blood type, component ISBT code and description, storage method and additional characteristics. The EIN of the example product is found on the left: N001112172541. The meaning of each block that forms this EIN is described as: supplier code "N0011", year of donation "12" (short for 2012), and internal number "172541"

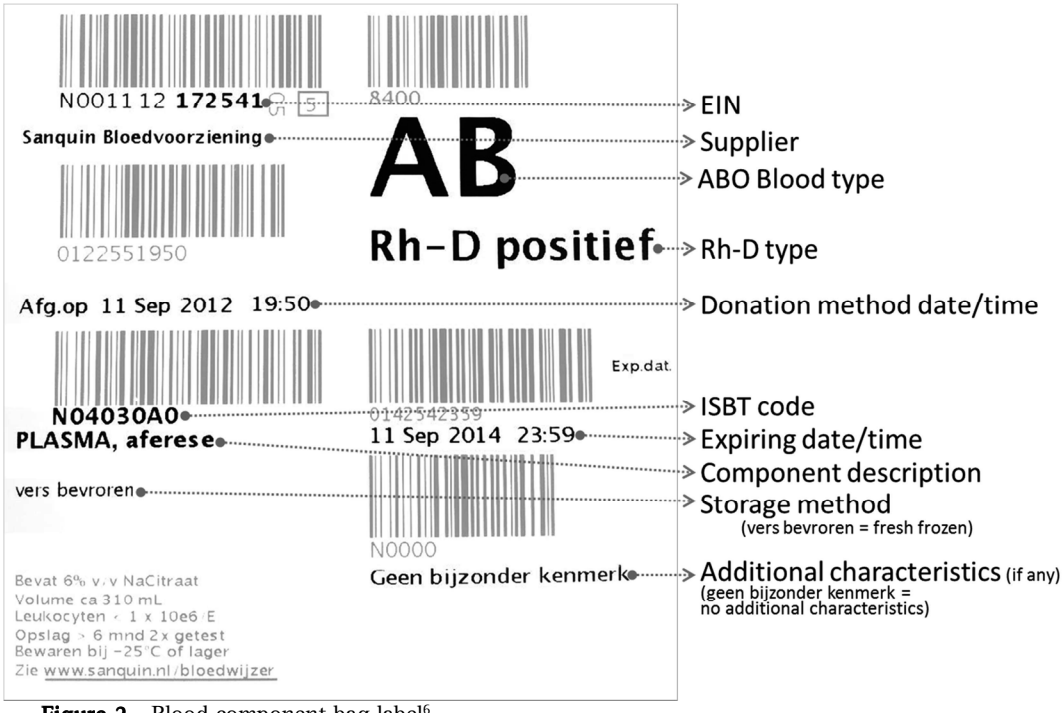


Figure 2 - Blood component bag label[6]

Some hospitals, due to historic or unknown reasons, split the EINs to store the information in their blood transfusion databases. Supplier, year and product internal number stored along with the component description (also recorded according to hospitals codes). Thus in these hospitals we could only retrieve the data in blocks and rebuild the EINs to merge them to the blood bank (donors) information. Some EINs could not be found in the blood bank database during the linkage because one or more blocks were wrongly recorded or missing.

A final example of an obstacle in the merging phases of our studies is about component codes. When databases from hospitals were retrieved the information on blood component description were included along with the information of patients and transfusions. Components were, not surprisingly, coded as hospitals use them in their routine, despite of the standardised International Society of Blood Transfusion (ISBT) product code which has been in use and accessible to hospitals for several years in the Netherlands.[7] The ISBT code can be seen in the component bag label (Figure 2) above the component description: N04030A0.

To increase reliability the data provided by the blood bank and by the hospitals was cross-matched. Based on component descriptions used in each hospital a table of possible and correct codes to match hospitals to the blood bank was created and used as the key for the crossmatch. It was observed, during the compilation of the component descriptions, that similar products are coded differently by different hospitals. Moreover, identical codes are used, in different hospitals, to denote different final components. For example, the code 'TC' is used by one hospitals for standard buffy coat derived platelets and in another for standard apheresis

platelets. A third hospital uses it to describe buffy coat and apheresis derived components that have been hyper-concentrated (i.e. medium removed). Other hospitals code platelets as 'PRP' or 'TCB' instead of 'TC'. There are also hospitals that use the same code for different components (e.g. buffy coat and apheresis) and their distinction could be only made via cross-match.

In the crossmatch some products did not meet the component description of the blood bank. For example, a hospital affirmed that the transfused product was a plasma component. However, in the blood bank database the component EIN referred to red blood cells.

Individually the merging problems described above can be solved manually. It is possible to check products one at the time in the hospital or blood bank systems to find the root of each individual error and correct it. However, in a data warehouse setting where individual traceability is not as important as in the cases of transfusion reactions, the adopted strategy was to identify and exclude from our analyses any patient who had one or more transfusions whose EIN was not found or did not crossmatch with the component description of the blood bank. These problems occurred in 0.3% of the blood components transfused; a very small percentage which will have limited effect on the validity of the study results. The exclusion of patients was to ensure reliability of results and estimations.

To avoid the described merging problems, hospitals should be encouraged to adapt their system to record information in the same pattern used by the blood supplier (i.e. blood bank). Currently a national blood transfusion data warehouse is being implemented in the Netherlands (DTD).[8] A standard for recording information such as ABO Rhesus D blood type, complete EIN codes (instead of the internal number) and the recommended use of ISBT codes can be developed and promoted when hospitals join the project.

Modelling

Modelling is a crucial step in the transformation of data to information. Its goal is to translate data into a single – or a few – equations or to estimate a complex, possibly non-linear and multi-correlated relationship. Modelling must be carefully planned and executed, especially in the selection of co-variables that affect the relationship between exposure and outcome. These variables, when masking the causal relationship between exposure and outcome, may lead to spurious associations, in which case they are called confounders.

Large datasets, in the presence of bias, are prone to p-value fallacy: misinterpretation of p-value and conclusion.[9] In large sample sizes p-values go quickly to zero, because p-value calculations are based on formulas that have sample sizes in their denominator. Thus, a naïve person may look only at the p-value significance from a study with big sample size and be trapped to make fast association conclusions. This happens because large samples are "too big to fail".[10] Analogously, confidence intervals tend to be narrow because they are also based on formulas that have sample sizes in their denominator.

Another challenge is to decide how to include variables in a model. The correct "shape" of the relationship between exposure, confounding variables and the outcome is often unknown. This decision is fundamental and often not straightforward.

When there is no confounding and only two variables are being modelled, a simple scatterplot of the variables gives the answer as

shown in figure 3. For the three hypotetical xy relationships shown in this figure the same polinomial equation can be used:

$$y=a+bx+cx^2+dx^3.$$

The values of a, b, c and d can be easily estimated by parametric methods such as maximum-likelihood estimations.[1] However, when the relationship between two variables are under the influence of several other variables the decision which model to use is far more complex. The underlying relationship (or trend) between variables can not be visualised by scatterplots and the decision how each term should be included in the model is, as matter of fact, a guess relying mostly on peoples' expertice.

One way of modelling a relationship is by using 'saturated models'. These models have as many estimated parameters as value levels. In a saturated model each variable and interaction is entered as an indicator variable. It means that each possible discrete value of each variable represents in the model one category. Although very useful, they generally leave too few degrees of freedom to estimate variability.[1] Thus, they can only be used in case of very big sample sizes, as in our study (**chapter 2**).

By including variables as categorical (i.e. indicator) variables in time varying cox regression models, we avoided the problem of guessing the "shape" of the relationship between exposure, confounding variables and the outcome. With that, we were likely to have achieved optimal fitting for the included variables.

Having selected the variables to be included in the model and how to include each one, we still had to judge whether the adjustment made was good enough. We still had to answer the question whether the chosen model subtracts from the exposure the effect of the confounding variables, i.e. does the model give an unbiased estimation?

The sad truth is, there is no way of judging whether a model is perfect. This is due to "unknown variables" which could lead to bias and are unmeasurable by definition (for being unknown). Thus, prior to modelling, a careful consideration about which variables to include is crucial. A useful tool in this process is the directed acyclic graph (DAG). A DAG is a diagram of causal pathways.[1] DAGs are useful to define and cover all possible and measurable variables and the relationships between those variables, outcomes and exposures.

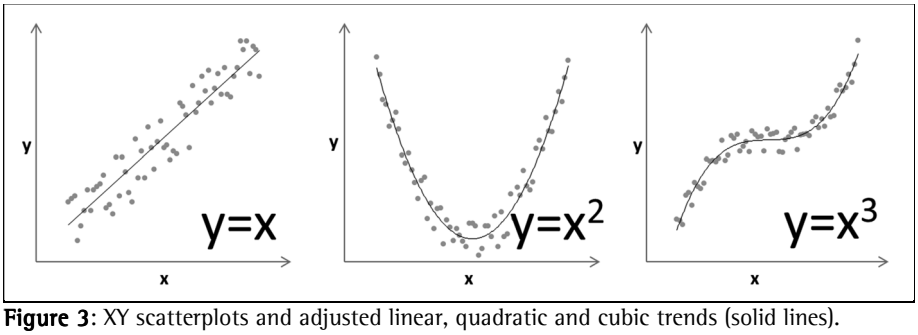


Figure 3: XY scatterplots and adjusted linear, quadratic and cubic trends (solid lines).

After modelling, sensitivity analyses and global goodness of fit tests can be carried out, as well as post-tests specific to each model (e.g. regression models or multi-level models). Additionally, we empirically observed that two opposite exposure categories have non inverse estimations when these estimations are biased. In contrast, unbiased estimations will show an inverse relationship for opposite exposure categories. This fact can be observed in Table 2. For biased (crude) estimations the estimation of male exposure is not the inverse of the opposite category (i.e. 1/female). Note that in this example crude estimations for both exposure categories have a hazard ratio above 1, thus both are representing risk in relation to the other. This is conceptually and mathematically impossible. After adjustments, using saturated models, hazard ratios of male and female showed a perfect inverse relationship:

male=1/female

(i.e. difference: male-1/female=0.00000).

In this way the opposite category works as a negative control to verify the models efficiency in correcting for the presence of confounding. In conclusion, an observational study, by collecting a large sample over 10 years made it possible, through saturated models, to have unbiased, inverse exposure estimations.

To be able to judge the validity of study results, high quality reporting is also needed. The next session discusses our findings about the use of the terms 'retrospective' and 'prospective' to describe observational studies and its relationship with quality of the report.

Table 2 – Hazard ratios for different exposure in a crude and adjusted model.

	Exposure		
Model	***Female***	***Male***	$Male - \frac{1}{\text{Female}}$
Crude hazard ratio	1.12612	1.40439	0.51639
Adjusted hazard ratio	1.02572	0.97492	0.00000

OBSERVATIONAL STUDIES REPORT

Past

We live in an era with high-speed science production and without knowledge boundaries. The new ways of learning go beyond books, guidelines and printed journals. The list includes Google, Wikipedia, PubMed, WebOfScience, GoogleBooks, Cochrane, UpToDate and many other medical specialised websites. People use all sorts of information sources to learn and update themselves.

In this new modern scenario authors are under pressure to prove not only their efforts to perform high quality research but also to write in a convincing and attractive way. Since evidence-based medicine was formulated and published in the early 90's[11,12] the supposed "hierarchy of clinical research" became widespread. As a consequence, authors of scientific papers were encouraged to put studies as high as possible in the pyramid of evidence.[1,13,14] The pyramid of evidence was previously presented in chapter 1, figure 6.

Observational studies (i.e. case control and cohort studies) are lower in the pyramid than experimental studies (i.e. randomised and non-randomised controlled trials). Observational studies often use secondary data (i.e. uncontrolled data generation) while experimental studies most of the time use primary data (i.e. controlled data generation). In trials a lack of bias (or low bias) is expected. Trials are designed to control and register any variation in outcome due to extraneous factors and account for those variations in comparisons across groups. However, in observational studies the chance of bias is bigger if variations due to extraneous factors surrounding the exposure-outcome relationship were not thoroughly recorded to later take into account when the groups are compared.[1]

Historically, writers used the term 'retrospective' as synonym of case-control studies and 'prospective' as synonym of cohort studies. As cohort studies are higher in the pyramid of evidence, they are closer to experimental studies and therefor studies labelled 'prospective' (i.e. cohort) were believed to be more reliable than studies labelled 'retrospective'. Later the terms 'retrospective' and 'prospective' lost their connections with the corresponding study designs and became adjectives constantly used despite of their various definitions.[1]

Over the years the terms became a trend that can be seen in the methods section of papers and also in titles and abstracts. The terms 'retrospective' and 'prospective' by themselves do not explain properly how the research was performed but lead naïve readers to believe that studies using the term 'prospective' are better than studies using 'retrospective' because 'prospective' could be interpreted as higher in the pyramid of evidence, resulting in bigger credibility. Besides that, there is no rational basis for this connotation, since there is no consensus about the meaning of these terms.

Several initiatives were taken by members of the scientific community to improve quality of research and reporting. Guidelines were produced and journals encourage authors to follow them.[3,15-22] One of these initiatives is the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)[16,17]. The STROBE, first published in 2007, recommended that authors refrain from the use of the terms 'retrospective' and 'prospective'. (Box 1)

Box 1 - STROBE recommendation about the terms 'retrospective' and 'prospective'

"We recommend that authors refrain from simply calling a study 'prospective' or 'retrospective' because these terms are ill defined. One usage sees cohort and prospective as synonymous and reserves the word retrospective for case e control studies. A second usage distinguishes prospective and retrospective cohort studies according to the timing of data collection relative to when the idea for the study was developed. A third usage distinguishes prospective and retrospective case e control studies depending on whether the data about the exposure of interest existed when cases were selected. Some advise against using these terms, or adopting the alternatives 'concurrent' and 'historical' for describing cohort studies. In STROBE, we do not use the words prospective and retrospective, nor alternatives such as concurrent and historical. We recommend that, whenever authors use these words, they define what they mean. Most importantly, we recommend that authors describe exactly how and when data collection took place."[16]

Present

Six years after the publication of the STROBE we systematically reviewed the use of the terms 'prospective' and 'retrospective' in observational studies published in top clinical journals. First we quantified to what extent authors follow the guideline's recommendations to not use the terms 'prospective' and retrospective. Then we looked if the use of these terms was associated with the quality of reporting measured as STROBE adherence score.

Analysing 150 observational papers in leading clinical journals we concluded that the STROBE's recommendation to refrain from using the terms 'prospective' and 'retrospective' has largely been ignored, also in journals recommending to use the STROBE guideline. Usage of the terms was not associated with the quality of the report. (**chapter 7**)

Future

Currently more than 130 journals recommend the STROBE guidelines. The STROBE explanation and elaboration, that advices about the non-use of the terms 'retrospective' or 'prospective' has being published once more, this time in Internal Journal of Surgery in 2014. However no changes were observed - also - in the recommendation of the non-use of the terms 'retrospective' and 'prospective'.

Figure 4 - shows, for the years of 2005 to 2015, the number of studies in PubMed with no restrictions on the study design, the number of epidemiological studies in PubMed and the percentage of papers that use the terms 'retrospective' or 'prospective' in the title or abstract in this two groups. This graph is an update of our publication that shows the same graph for the years 1975 to 2010. It can be observed that the percentage of papers that use the terms 'retrospective' or 'prospective' continues the increasing trend that was already observed previously.

On average more than 6% of the papers use the terms 'retrospective' or 'prospective' (reference 2005 to 2015). There is no doubt that the terms are routinely used and journals continue to publish papers regardless of their use. In fact, it is difficult to think of another adjective that is used with the same frequency.

We reaffirm our paper's vision of the importance of a discussion, on the merits and perils of using these terms, among a broad representation of the medical scientific community. The outcome should be, either making people aware and following the advice to avoid the terms 'prospective' and retrospective, or to update the STROBE with the subtract of the recommendation along with an expert definition of the terms.

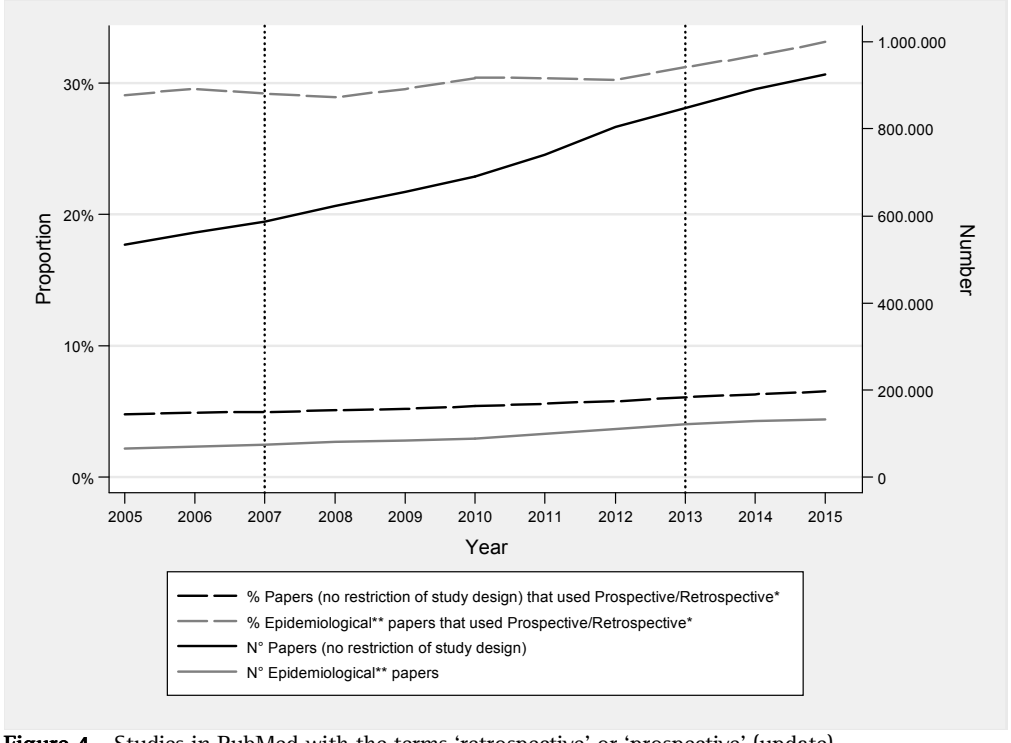


Figure 4 - Studies in PubMed with the terms 'retrospective' or 'prospective' (update)

* Epidemiologic Studies[Mesh] NOT Seroepidemiologic Studies[Mesh]
** Abstract or Title
All searches were restricted to papers with abstract available
Dashed lines refer to proportions and are showed in the primary y axis (left side)
Continuous lines refer to absolute numbers and correspond to the secondary y axis (right side)
Black lines have no restriction of the study design
Grey lines have restriction of study design (epidemiological papers excluded seroepidemiologic studies)
The vertical dotted lines corresponds to 2007: year of the STROBE guidelines publication and
2013: year of reference of our systematic review

FEMALE DONORS AND THEIR PREGNANCY HISTORY

Past

Discordance between the sex of the blood donor and the sex of the recipient (patient) has been described since the 40's. One of the first reports describing this relationship studied the exposure to male versus female and the exposure to sex-match versus sex-mismatch in 2,720 patients. This paper concluded that receiving blood from a female donor would increase the incidence of transfusion reactions compared to male donors.[23] One year later an American team reported the survival rate between 205 babies with erythroblastosis who were treated by exchange transfusions.[24] The authors conclude that blood from female donors was beneficial. However, they remarked that the female donors in their study were all heathy, young, non-pregnant and with no recent history of delivery.[24]

In response to the American team, a Canadian team published a study with 74 similar profile patients who went through the same treatment (exchange transfusions).[25] This paper, besides survival, also included severity of disease (described as red blood cells count) as confounding variable and concludes that there was no advantage of using female blood in the transfusion therapy.[25] Those publications were however from the time when whole blood was the only product available for routine transfusion. In the 70's, with the implementtation of centrifuge techniques to spin-separate whole blood into plasma, platelets and red blood cells, the whole blood therapy gave way to components therapy until, in the 80's, the latter became the standard care.[26,27]

When whole blood was divided into components, it was the plasma that inherited the association between donor's sex and side effects. In fact, plasma from female donors was reported to increase the risk of transfusion-related acute lung injury (TRALI) in several studies.

TRALI has been described to be caused by donor-derived leukocyte antibodies, which can be directed either against the human leukocyte antigens (HLAs) or against the human neutrophil antigens (HNAs). HLA and HNA antibodies occur most frequently in plasma-rich components from parous women.[28] Consequently, several countries adopted a policy of male-only plasma or of male-predominant-plasma to prevent TRALI.[29-32] In 2015 it was estimated that more than 800 million people in 17 countries were living under either of these policies.[30] The male-only plasma or male-predominant-plasma strategies proved to be efficient and the number of TRALI cases has fallen significantly since then.[28-32]

Meanwhile, red blood cells were forsaken. It was only in 2011 that a Dutch study reported an increased hazard ratio for mortality after sex-mismatched transfusions, compared to sex-matched transfusions in a cohort of 11,211 patients who received 96,009 blood components (73,293 of them red blood cells).[33] Table 3 shows the hazard ratio of receiving sex-mismatched blood components compared to sex-matched components in this study. The hazard ratio was 2.1 (95% confidence interval, CI: 1.2 to 3.6) for transfusions of any blood component (platelets, plasma and red blood cells) and 2.4 (CI: 1.1 to 5.2) for transfusions of red blood cells.

Table 3 – Hazard Ratio of sex-mismatched compared to sex-matched transfusions.[33]

Component	Hazard Ratio*
Platelets, plasma and red blood cells	2.1 (1.2 to 3.6)
Red blood cell	2.4 (1.1 to 5.2)

* Hazard Ratio and (95% confidence interval), half year follow up for patients between 1 and 55 years old.

These results kindled interest in the relationship between the sex of (red blood cells) donors and mortality. Subsequent publications looked at this relationship in two ways:

1. Female exposure: receiving red blood cells from female donors compared to receiving red blood cells from male donors, regardless of the patients being female or male.[34]
2. Sex-mismatch exposure: receiving red blood cells from a donor whose sex is the opposite of the sex of the patient. Namely, female patients transfused with red blood cells from males donors and male patients transfused with red blood cells from female donors.[35-38]

All recent studies that reported an effect on mortality, reported a stronger effect in male patients than in female patients and stronger in the exposure to female red blood cells than to male red blood cells. In conclusion, an increased mortality after sex mismatch transfusions for male patients (i.e. exposed to female red blood cells).

Present

In a cohort of 60,912 patients who received a total of 230,099 red blood cells transfusions in 6 major hospitals in the Netherlands (3 university hospitals) we studied the effect of donor sex on the mortality of patients (**chapter 2**). Beyond the previously published papers we included the information of prior pregnancy of female donors. Our hypothesis was that one possible explanation for the observed association could be an immunological mechanism based on immunological changes occurring in female donors during pregnancy.

In our study we show the association between recipient survival and red blood cell transfusion from female donors, with and without a history of pregnancy, to depend on a positive pregnancy history of those female donors and to be specific for male recipients. Especially male recipients under 50 years of age. We also meta-analysed our results with the recently published literature papers. This analysis showed that our results are in line with different studies on the association of donor sex and recipient survival (Figure 5).

Pooling the results of 6 different hospitals incorporates concepts of the Data-Information-Knowledge Hierarchy, including the associated problems and solutions. Two key points had a special place in our study (i) merging different source databases and (ii) modelling. These two topics were already discussed previously in this chapter under the topic “validity”.

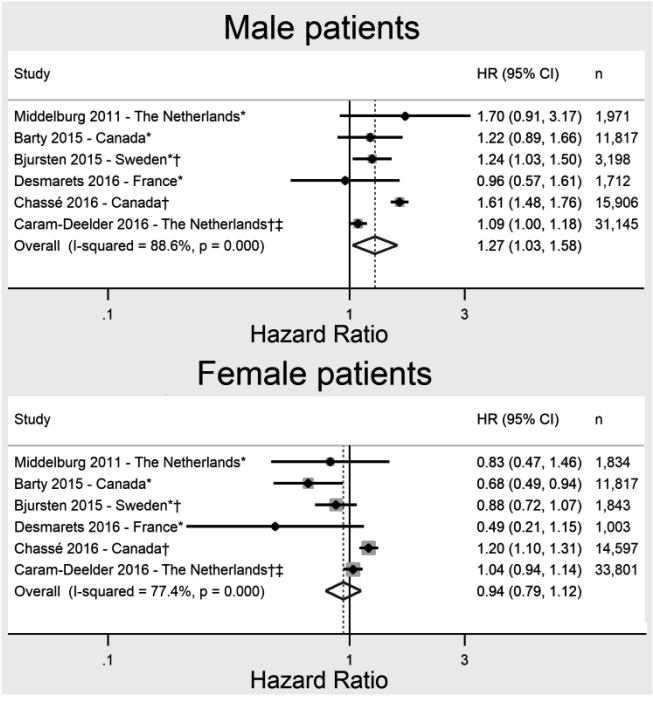


Figure 5 – Recently published papers on the association of recipient mortality female blood donors

HR (95%CI): 95% confidence interval of the hazard ratio (HR). Weights are from random effects analysis
* Female exposure was recalculated from sex-mismatched transfusions. † Hazard ratio per transfusion powered to the mean number of transfusions in that study. ‡ Regardless of pregnancy

Future

Red blood cells are nowadays given, in the Netherlands, regardless of the sex of the donor. In fact the sex of the donor is not written on the blood bag and can only be retrieved via the blood bank. This means that in practice the sex of the donor is randomly assigned to patients, and both patients and health professionals are unaware of it.

We provided a strong piece of evidence to support either a change in recommendation from blood transfusion guidelines or at least to encourage further studies focusing on pregnancy history of female donors and its effect on male and female patients, children and adult patient groups.

Further studies focusing on pregnancy history and the biological mechanism of how blood from females after pregnancy affects survival would reveal important information. A change in the guidelines would involve changes in the blood bag labels and in hospital routines of transfusing female or male blood, depending on the profile of the patient. Such changes would affect the whole transfusion chain. This decision goes beyond this doctoral thesis but we believe it is important that policy makers are aware of our findings. Our results call for at the least an immediate improvement regarding to the recording of pregnancy history of the female donors. Currently female donors are asked if they “Have ever been pregnant? ” at the time of the first donation and “have you been pregnant since the last donation” at all subsequent donations. This way of recording data proved to be not efficient and resulted in an important limitation of our study. Recording “date of the first known pregnancy” and “date of the most recent known pregnancy” in all donations (instead of yes/no) would already provide the opportunity of more powerful future studies.

PLATELETS STORAGE TIME

Past

The history of the use of platelets goes back to the 50's. Children with leukaemia who were bleeding due to chemotherapy were treated with platelets in an experimental way.[39] In the 70's this was already a routine therapy. In the 80's technologies such as removing leukocytes from platelets became available and made possible to store platelets initially for a maximum of 72 hours. By the year 2000, after 20 years of studies, production and material improvements, fresh filtered platelets were routinely stored for 5 days in the big majority of blood banks in Western countries.[40] With the developments of platelet additive solution PAS-C, platelets could be finally stored up to 7 days.[41,42]

Nowadays platelet components can be produced in three ways as shown in the Figure 6. The first method is called platelet rich plasma (PRP): from whole blood donations, red blood cells are separated from plasma and platelets via soft spin. Then plasma is separated from platelets via hard spin. The buffy coats are then pooled. No spin is needed after pooling. This component is called "PRP platelets in plasma". The second method is called "buffy coat method": red blood cells and plasma are separated from buffy coats (containing leukocytes and platelets) via hard spin. Then buffy coats and either additional plasma from one of the original whole blood donations or platelets addictive solution (PAS) are pooled and mixed via a soft spin. Components produced by this method are called "buffy coat platelets in plasma" or "buffy coat platelets in PAS" depending on the medium used. The component "Apheresis platelets in PAS or plasma" come from a single donations unit of platelets via apheresis machines.[43]

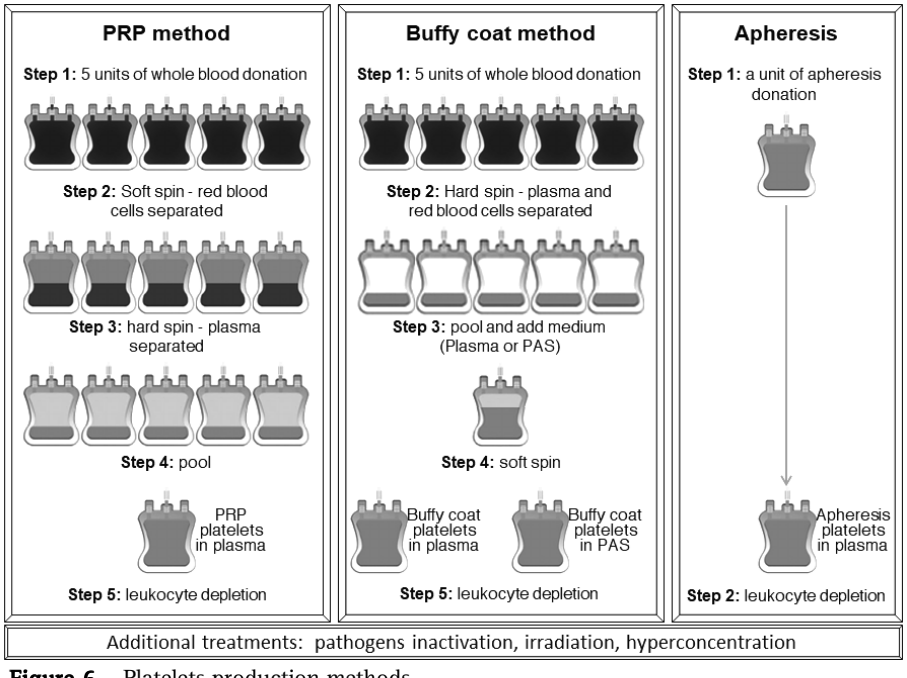


Figure 6 – Platelets production methods

In the Netherlands platelets are produced by apheresis and buffy coat method. They are always leukoreduced but irradiated and hyper-concentrated only on indication. Pathogen inactivation is not a standard product.[27,44]

Endogenous platelets circulate with an average lifespan of 5–9 days in humans. Thus, the body must generate and clear 10% of platelets daily to maintain normal physiological blood platelet counts.[45] It means that when blood is drained from donors to the bag, it naturally contains 'new' and 'old' platelets (i.e. cells produced up to 9 days before). On average platelets in the bag are 5 days old. These cells will age while waiting inside the blood bag to be transfused.

In addition, there are deleterious changes in the platelet structure and function during component production and storage, namely storage lesion. This will further diminish the number of cells (platelets count) and the efficiency of platelet components when transfused.[46] Because platelets are stored at room temperature they are more prone to bacterial contamination than other blood components that are stored at lower temperatures. In the Netherlands, each platelet unit is screened immediately after preparation for aerobic and anaerobic bacterial contamination using the BacT/ALERT system.[27]

As with any other blood component, the decision of transfusing should be a balance between the benefits, the risk of side effects and costs.[47] Storage time is an important variable in this context. Storage time has been reported to be (negatively) associated to several outcomes.

Present

Two complementary meta-analyses (**chapters 4 and 5**) were performed to investigate the association between storage time of platelets and clinical outcomes, including platelet counts and derived measures.

We systematically reviewed 4,234 abstracts and titles in 7 publication search engines. Out of those, 59 publications had one or more outcomes meta-analysed. Table 4 gives a combined overview of all meta-analysed outcomes. We observed that fresh platelets were superior to old platelets for several outcomes: all platelet measurements, transfusion reactions (if platelets were non-leukoreduced), and transfusion interval. It also showed a possible superiority of fresh over old platelets for the numbers of platelet transfusions in haematological patients.

Several research questions were raised from the results of the meta-analyses, of which two were studied further. The first was the diagnosis of platelet refractoriness – i.e. failure to achieve the desired platelet count following a platelet transfusion – after transfusions of fresh and old platelets (**chapter 6**). We found that the effect of storage time was remarkably stable across different cut-off values for successful transfusions. For 24 corrected increment count, irrespective of the cut-off used, the number needed to treat was 18, to prevent one failed transfusion because of an old product where a fresh product would have been successful.

This study by using the meta-analysis-derived distribution of the 1h and 24h platelet count increments and corrected increment for fresh and old platelets embodies the concepts of "Data–Information–Knowledge Hierarchy" transforming data into knowledge as described in the chapter 1.

In the second study we addressed the association between storage time of platelet concentrates, stored for up to seven days, and the interval to the next prophylactic platelet transfusion (**chapter 3**). An algorithm to reveal patterns and trends of the platelets transfusion was developed. This algorithm was validated and used to select periods of platelet transfusion dependency. In this study 10 different hospital databases, from two sources, were merged. Merging and modelling challenges faced in this study were discussed previously in this chapter under the topic "validity". With 94% specificity the algorithm selected 4,441 hemato-oncology patients, who had received 12,724 transfusions, from a cohort of 29,761 patients, who received 140,896 platelet transfusions. In line with the meta-analysis findings an association was shown, in the selected cohort, between increased storage time and decreased transfusion interval for all buffy coat-derived platelet components but not for apheresis components. It was also shown that, in spite of the association observed for buffy coat-derived products, the additional outdating associated with 5-day storage of platelets would easily outweigh the potential benefit resulting in 8.6% increase in platelet components waste.

In conclusion, we have shown in chapters 3 to 6 that old platelets were inferior to fresh platelets for all measurements, transfusion interval and the need of additional platelet transfusions. Superiority of fresh platelets in transfusion reactions was not observed when components were leuko-reduced.

Table 4 – combined meta-analyses results

Outcome	Storage time association*
Measurement - Count increment 1h	**Favours fresh** | No difference | Favours old
Measurement - Count increment 24h	**Favours fresh** | No difference | Favours old
Measurement - Correct count increment 1h	**Favours fresh** | No difference | Favours old
Measurement - Correct count increment 24h	**Favours fresh** | No difference | Favours old
Measurement - Recovery	**Favours fresh** | No difference | Favours old
Measurement - Survival	**Favours fresh** | No difference | Favours old
Clinical - Safety: Transfusion reactions†, before universal leukoreduction	**Favours fresh** | No difference | Favours old
Clinical - Safety: Transfusion reactions†, after universal leukoreduction	Favours fresh | **No difference** | Favours old
Clinical - Safety: Complications‡	Favours fresh | **No difference** | Favours old
Clinical - Safety: Mortality	Favours fresh | **No difference** | Favours old
Clinical - Efficacy: Transfusion interval	**Favours fresh** | No difference | Favours old
Clinical - Efficacy: Bleeding	Favours fresh | **No difference** | Favours old
Clinical - Efficacy: Transfusion need, platelets	Favours fresh | **No difference** | Favours old
Clinical - Efficacy: Transfusion need, red blood cells	Favours fresh | **No difference** | Favours old
Clinical - Efficacy: Transfusion need, plasma	Favours fresh | **No difference** | Favours old

* Favours fresh: favours platelets stored shorter than the contrast category (old platelets); No difference: fresh equals to old platelets; Favours old: favours platelets stored longer than the contrast category (fresh platelets).
† transmission infection, allogenic transfusion reaction, febrile non haemolytic transfusion reaction (FNHTR) and transfusion related acute lung injury (TRALI);
‡ infections and overall complications

Future

Blood banks and researchers are continuously working to make blood component transfusions more efficient, safer and cost-effective. For platelet components the next steps are the evaluation of pathogen reduction and the possibility to store platelets bellow room temperature. Storage at lower temperatures has shown potential to improve efficiency in a situation of acute haemorrhage[44,48]

Pathogen reduction could potentially enhance safety by reducing bacterial and viral contamination and also reduce alloimmunization in multiply transfused patients.[44,48] Pathogen reduction can also support the extension of shelf-life of platelet components beyond the current 7 days allowed. The storage bags and storage conditions currently in use, however, reduce platelets efficiency over time, as shown in this thesis and several other studies. Hence, new storage bag materials and coatings that promote platelet respiration, while not supporting platelet adhesion or activation are needed. These materials are in development and tests are promising.[49] Nevertheless extended shelf-life is only feasible if platelets can be temporarily inactivated while stored, due to their natural lifespan of 5 to 9 days.[45,46,49]

A randomised, single-blinded, multicentre controlled trial was performed in the Netherlands, Norway and Canada. The "Pathogen Reduction Evaluation and Predictive Analytical Rating Score" (PREPAReS) trial compared standard plasma-stored platelets to platelets pathogen reduced via the Mirasol system. In both arms platelets were allowed to be stored for up to 7 days. The primary endpoint of the trial was WHO grade ≥2 bleeding complications.[44] PREPAReS' patients inclusion was recently closed and results are expected.

REFERENCES

1. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
2. Uman LS. Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent*. 2011;20(1):57-59.
3. Liberati A, Altman DG, Tetzlaff J, et al. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Annals of internal medicine*. 2009;151(4):W-65-W-94.
4. Johnson BT. Meta-Analysis: The Original 'Big Data'. 2013; http://meta-analysis.ning.com/profiles/blogs/meta-analysis-the-original-big-data. Accessed 06/01/2017.
5. Sauerland S, Seiler CM. Role of systematic reviews and meta-analysis in evidence-based medicine. *World journal of surgery*. 2005;29(5):582-587.
6. Tynngard N, Boknas N, Trinks M, Dreimane A, Berlin G. Comparison of fresh and stored platelets, obtained from a single donor and donation, transfused to a single patient. *Transfusion*. 2015;Conference(var.pagings):24A-25A.
7. ISBT. ST-001 ISBT 128 Standard Technical Specification-v5.7.1. https://www.iccbba.org/tech-library/iccbba-documents/technical-specification (accessed March 2017).
8. van Hoeven LR, Hooftman BH, Janssen MP, et al. Protocol for a national blood transfusion data warehouse from donor to recipient. *BMJ open*. 2016;6(8):e010962.
9. Goodman SN. Toward evidence-based medical statistics. 1: The P value fallacy. *Annals of internal medicine*. 1999;130(12):995-1004.
10. Lin M, Jr HCL, Shmueli G. Research Commentary—Too Big to Fail: Large Samples and the p-Value Problem. *Information Systems Research*. 2013;24(4):906-917.
11. Group E-BMW. Evidence-based medicine. A new approach to teaching the practice of medicine. *Jama*. 1992;268(17):2420-2425.
12. Smith R, Rennie D. Evidence-based medicine--an oral history. *Jama*. 2014;311(4):365-367.
13. Rosner AL. Evidence-based medicine: revisiting the pyramid of priorities. *Journal of bodywork and movement therapies*. 2012;16(1):42-49.
14. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. 2016;21(4):125-127.
15. American Society of Anesthesiologists Task Force on Perioperative Blood T, Adjuvant T. Practice guidelines for perioperative blood transfusion and adjuvant therapies: an updated report by the American Society of Anesthesiologists Task Force on Perioperative Blood Transfusion and Adjuvant Therapies. *Anesthesiology*. 2006;105(1):198-208.
16. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *International journal of surgery (London, England)*. 2014;12(12):1500-1524.
17. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *International journal of surgery (London, England)*. 2014;12(12):1495-1499.
18. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ (Clinical research ed)*. 2009;339.
19. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ (Clinical research ed)*. 1996;313(7057):570-571.
20. Brand RA. Standards of reporting: the CONSORT, QUORUM, and STROBE guidelines. *Clinical orthopaedics and related research*. 2009;467(6):1393-1394.
21. Simon AK, Rao A, Rajesh G, Shenoy R, Pai MB. Reporting guidelines in health research: A review. *Indian journal of public health*. 2015;59(2):122-128.
22. Vandenbroucke JP. STREGA, STROBE, STARD, SQUIRE, MOOSE, PRISMA, GNOSIS, TREND, ORION, COREQ, QUOROM, REMARK... and CONSORT: for whom does the guideline toll? *Journal of clinical epidemiology*. 2009;62(6):594-596.
23. Anderson CD, Seldon TH. Influence of sex of donors and recipients on transfusion reactions. *Proceedings of the Staff Meetings of the Mayo Clinic*. 1948;23(6):52, 1948.
24. Allen FH, Jr., Diamond LK, Watrous JB, Jr. Erythroblastosis fetalis: the value of blood from female donors for exchange transfusion. *The New England journal of medicine*. 1949;241(21):799-806.
25. Sacks MS, Spurling CL, Bross ID, Jahn EF. Study of the influence of sex of donor on the survival of erythroblastotic infants treated by exchange transfusion. *Pediatrics*. 1950;6(5):7, 1950.
26. Zielinski MD, Jenkins DH, Hughes JD, Badjie KS, Stubbs JR. Back to the future: the renaissance of whole-blood transfusions for massively hemorrhaging patients. *Surgery*. 2014:155(5):883-886.
27. de Vries R, Haas F. English translation of the Dutch Blood Transfusion guideline 2011. *Vox Sanguinis*. 2012;103(4):363.
28. Middelburg RA, van Stein D, Briet E, van der Bom JG. The role of donor antibodies in the pathogenesis of transfusion-related acute lung injury: a systematic review. *Transfusion*. 2008;48(10):2167-2176.
29. Wiersum-Osselton JC, Middelburg RA, Beckers EA, et al. Male-only fresh-frozen plasma for transfusion-related acute lung injury prevention: before-and-after comparative cohort study. *Transfusion*. 2011;51(6):1278-1283.
30. Schmickl CN, Mastrobuoni S, Filippidis FT, et al. Male-predominant plasma transfusion strategy for preventing transfusion-related acute lung injury: a systematic review. *Critical care medicine*. 2015;43(1):205-225.
31. Chapman CE, Stainsby D, Jones H, et al. Ten years of hemovigilance reports of transfusion-related acute lung injury in the United Kingdom and the impact of preferential use of male donor plasma. *Transfusion*. 2009;49(3):440-452.
32. Gajic O, Yilmaz M, Iscimen R, et al. Transfusion from male-only versus female donors in critically ill recipients of high plasma volume components. *Critical care medicine*. 2007;35(7):1645-1648.
33. Middelburg RA, Briet E, van der Bom JG. Mortality after transfusions, relation to donor sex. *Vox Sang*. 2011;101(3):221-229.
34. Chasse M, Tinmouth A, English SW, et al. Association of Blood Donor Age and Sex With Recipient Survival After Red Blood Cell Transfusion. *JAMA internal medicine*. 2016.
35. Barty RL, Cook RJ, Liu Y, Acker J, Eikelboom JW, Heddle NM. Exploratory analysis of the association between donor sex and in hospital mortality in transfusion recipients. *Transfusion*. 2015;Conference: AABB Annual Meeting 2015 Anaheim:September 2015.
36. Bjursten H, Dardashti A, Bjork J, Wierup P, Algotsson L, Ederoth P. Transfusion of sex-mismatched and non-leukocyte-depleted red blood cells in cardiac surgery increases mortality. *The Journal of thoracic and cardiovascular surgery*. 2015.
37. Desmarets M, Bardiaux L, Benzenine E, et al. Effect of storage time and donor sex of transfused red blood cells on 1-year survival in patients undergoing cardiac surgery: an observational study. *Transfusion*. 2016.
38. Holzmann MJ, Sartipy U, Olsson ML, Dickman P, Edgren G. Sex-Discordant Blood Transfusions and Survival After Cardiac Surgery: A Nationwide Cohort Study. *Circulation*. 2016;134(21):1692-1694.
39. Freireich EJ. Origins of platelet transfusion therapy. *Transfusion medicine reviews*. 2011;25(3):252-256.
40. Brand A, Novotny V, Tomson B. Platelet Transfusion Therapy: From 1973 to 2005. *Human Immunology*. 2006;67(6):413-418.
41. van der Meer PF. PAS or plasma for storage of platelets? A concise review. *Transfusion medicine*. 2016;26(5):339-342.
42. Ashford P, Gulliksson H, Georgsen J, Distler P. Standard terminology for platelet additive solutions. *Vox Sang*. 2010;98(4):577-578.
43. Stroncek DF, Rebulla P. Platelet transfusions. *Lancet (London, England)*. 2007;370(9585):427-438.
44. Ypma PF, van der Meer PF, Heddle NM, et al. A study protocol for a randomised controlled trial evaluating clinical effects of platelet transfusion products: the Pathogen Reduction Evaluation and Predictive Analytical Rating Score (PREPAReS) trial. *BMJ open*. 2016;6(1):e010156.
45. Grozovsky R, Hoffmeister KM, Falet H. Novel clearance mechanisms of platelets. *Current opinion in hematology*. 2010;17(6):585-589.
46. Mittal K, Kaur R. Platelet storage lesion: An update. *Asian journal of transfusion science*. 2015;9(1):1-3.
47. Eichbaum Q, Smid WM, Crookes R, et al. Apheresis in developing countries around the World. *Journal of clinical apheresis*. 2015;30(4):238-246.
48. Ness PM. Platelet transfusions: are we ready to chill out? *British journal of haematology*. 2017;178(1):7-8.
49. Farrugia BL, Chandrasekar K, Johnson L, et al. Perspectives on the use of biomaterials to store platelets for transfusion. *Biointerphases*. 2016;11(2):029701.

LIST OF PUBLICATIONS

In this thesis

1. Caram-Deelder C, Kreuger AL, Evers D, et al. Association of blood transfusion from female donors with and without a history of pregnancy with mortality among male and female transfusion recipients. *JAMA.* 2017;318(15):1471-1478.
2. Caram-Deelder C, van der Bom JG, Putter H, et al. Age of platelet concentrates and time to the next transfusion. *Transfusion;* In press.
3. Caram-Deelder C, Kreuger AL, Jacobse J, van der Bom JG, Middelburg RA. Effect of platelet storage time on platelet measurements: a systematic review and meta-analyses. *Vox sanguinis.* 2016;111(4):374-382.
4. Kreuger AL, Caram-Deelder C, Jacobse J, Kerkhoffs JL, van der Bom JG, Middelburg RA. Effect of storage time of platelet products on clinical outcomes after transfusion: a systematic review and meta-analyses. *Vox sanguinis.* 2017;112(4):291-300.
5. Caram-Deelder C, Kreuger AL, Rosendaal FR, van der Bom JG, Middelburg RA. Continuing use of the terms prospective and retrospective and quality of reporting of observational studies: time to update the STROBE guideline? *Int J Epidemiol.* 2016;45(2):587-589.
6. Caram-Deelder C, van der Bom JG, Kerkhoffs J-L, Middelburg RA. Storage time of platelet concentrates and the diagnosis of platelet refractoriness. (submitted).

Other publications

7. Braekkan SK, Caram-Deelder C, Siegerink B, et al. Statin use and risk of recurrent venous thrombosis: results from the MEGA follow-up study. *Research and Practice in Thrombosis and Haemostasis.* 2017;1(1):112-119.
8. Soares TH, de Bastos M, de Carvalho BV, et al. Prognostic value of computed tomographic pulmonary angiography and the pulmonary embolism severity index in patients with acute pulmonary embolism. *Blood Coagul Fibrinolysis.* 2013;24(1):64-70.
9. Pio SF, Ozelo MC, Dos Santos A, et al. Factor VIII inhibitors in patients with congenital severe haemophilia A and its relation to genotype. *Haemophilia.* 2012;18(6):e411-414.
10. Caram C, de Souza RG, de Sousa JC, et al. The long-term course of factor VIII inhibitors in patients with congenital haemophilia A without immune tolerance induction. *Thromb Haemost.* 2011;105(1):59-65.
11. Caram C, Monteiro-de-Castro MS, Caiaffa WT, et al. [Spatial-temporal distribution of potential blood donors at the Hemominas Foundation, Belo Horizonte, Minas Gerais State, Brazil, in 1994 and 2004]. *Cad Saude Publica.* 2010;26(2):229-239.
12. Rezende SM, Pinheiro K, Caram C, Genovez G, Barca D. Registry of inherited coagulopathies in Brazil: first report. *Haemophilia.* 2009;15(1):142-149.
13. Rezende SM, Caram C, Barca D, Pineiro K. *Perfil das coagulopatias hereditárias no Brasil 2007.* Ministério da Saúde Brasil (Brazilian National Health Department); 2008.
14. Galvao MA, Lamounier JA, Bonomo E, et al. [Emerging and reemerging rickettsiosis in an endemic area of Minas Gerais State, Brazil]. *Cad Saude Publica.* 2002;18(6):1593-1597.

ENGLISH SUMMARY

This thesis explores the potential of secondary data to answer questions regarding safety and efficacy of blood products. Secondary data is data that was generated for a different purpose than the research itself. In the first chapter an overview of the positive and negative effects of blood transfusion is given. In the following chapters secondary data is used for original research, combining databases of several hospitals, and in meta-analyses, combining results of different studies. In chapter 2 the relationship between patient survival and the sex of red cells donors was studied, while taking into account the pregnancy history of female donors and the sex of the patient. The results indicate that receiving a red cell transfusion from a female donor who has been pregnant is associated with increased mortality among male patients under the age of 50. In chapter 3 the relationship between 'time to the next platelet transfusion' and 'storage time of transfused platelet' was investigated. Results show that the interval between transfusions decreases as the age of buffy coat-derived platelet increases. This means that on a population level, more transfusions are needed when older platelets are transfused. However, this increase in the number of transfused units of platelets is unlikely to outweigh the benefit of reduced outdating and wastage, known to be associated with extended storage times. Chapters 4 and 5 describe two complementary systematic reviews and meta-analyses, addressing the relation between platelet storage time and clinical and laboratory measurement outcomes. Old platelets were inferior to fresh platelets for all platelet related counts, time between transfusions and the need of additional platelet transfusions. Fresh platelets were also associated with fewer transfusion reactions, but this association was absent when components were leuko-reduced. The results of these meta-analyses were used, in chapter 6, to estimate the underlying distribution of platelet increments and the probability of failure to reach the adequate count increment after fresh and old platelets transfusions. Inadequate increments can be used to diagnose transfusion failure or refractoriness. Chapter 7 addressed the methodological issue of the use of the terms 'prospective' and 'retrospective' in clinical observational research and its relationship with quality of reporting. These terms are still broadly used in spite of the recommendation to refrain from using them by guidelines and journals' editorial boards. The usage of the terms was, however, not associated with the quality of the report. Finally in chapter 8 the findings and concerns about validity of secondary data are discussed.

NEDERLANDSE SAMENVATTING

Dit proefschrift exploreert de mogelijkheden van secundaire gegevens bij het beantwoorden van vragen met betrekking tot de veiligheid en effectiviteit van bloedproducten. Secundaire gegevens zijn gegevens die gegeneerd zijn voor een ander doel dan het onderzoek zelf. In het eerste hoofdstuk wordt een overzicht gegeven van de positieve en negatieve effecten van bloedtransfusie. In de volgende hoofdstukken worden secundaire gegevens gebruikt voor origineel onderzoek, door het combineren van databases van diverse ziekenhuizen, en in meta-analyses, door het combineren van resultaten van diverse studies. Hoofdstuk 2 beschrijft een studie naar de associatie tussen de overleving van de patiënt en het geslacht van de donor van rode bloedcellen. Hierbij werd rekening gehouden met het geslacht van de patiënt en een eventuele zwangerschap in de voorgeschiedenis van vrouwelijke donoren. Voor mannen onder de 50 jaar bleek het ontvangen van een rode bloedceltransfusie van een vrouwelijke donor die ooit zwanger was geweest geassocieerd met een verhoogd overlijdensrisico. In hoofdstuk 3 werd de associatie onderzocht tussen 'tijd tot de volgende trombocyten transfusie' en 'bewaarduur van getransfundeerde trombocyten'. Het interval tussen twee opvolgende transfusies nam af bij het toenemen van de bewaarduur van trombocyten gemaakt van buffy coats. Dit impliceert dat op populatieniveau meer transfusies nodig zijn als oudere producten worden getransfundeerd. Deze toename in het aantal transfusies weegt echter niet op tegen het voordeel van verminderde verspilling door verlopen van producten bij een toename van bewaarduur. Hoofdstuk 4 en 5 beschrijven twee complementaire systematische reviews en meta-analyses over de associatie van bewaarduur van trombocytenconcentraten met diverse uitkomsten. Oudere trombocyten-concentraten bleken slechter te scoren dan verse producten op count increment en andere tellingen van trombocyten, het transfusie interval en het totale aantal benodigde transfusies. Verse trombocyten waren tevens geassocieerd met minder transfusiereacties, maar dit effect gold alleen voor niet-leukogereduceerde producten. De resultaten van deze meta-analyses werden gebruikt in hoofdstuk 6 bij het schatten van de onderliggende distributie van incrementen en de kans op het niet bereiken van een adequate opbrengst na transfusie van verse of oude trombocytenconcentraten. Bij herhaaldelijk lage incrementen is sprake van transfusiefalen of refractairiteit. Hoofdstuk 7 beschrijft het methodologische aspect van het gebruik van de termen 'prospectief' en 'retrospectief' in klinisch observationeel onderzoek gerelateerd aan de kwaliteit van de rapportage. Ondanks aanbevelingen van richtlijnen en redacties van tijdschriften om deze termen niet meer te gebruiken, worden deze termen nog altijd veel toegepast. Het gebruik van deze termen was echter niet geassocieerd met de kwaliteit van de rapportage. Tot slot wordt in hoofdstuk 8 de resultaten en zorgen omtrent de validiteit van secundaire gegevens bediscussieerd.

ACKNOWLEDGEMENTS

To Anske, for the trust and for giving me the opportunity of being part of your team. For never allowing me to give up, even when I wished you did. Thank you for the privilege of having you in my life as a role model of a woman, mother, boss, teacher and human being. You are one of the most inspiring people I have ever met.

To Rutger, for exposing me to your ginger brain! You were my best work mate and my biggest nightmare. You pushed me to the limit and then you encouraged me to go forward. I am grateful for that. Thank you for your enthusiasm, for dreaming big and motivating me to do the same. My achievements are also yours. I am proud for being your first PhD student.

To the committee and my paranymphs, I will always remember you for being part of one of the most important days of my life. It means a lot to me! I really appreciate your personal journeys and your commitment to science and teaching. To Dr. Suely for opening the doors of the research in Europe to me.

To every single person who contributed to my doctorate, directly or indirectly. To my Sanquin and Epi-LUMC colleagues, thank you for helping me to transform recorded data in knowledge. I do believe that life expectancy and quality of life increased considerably in the last century due to people like us. Thank you all, for each contribution, big or small.
"A dream you dream alone is only a dream. A dream you dream together is reality." [John Lennon]

To my parents, Regina and Julio; to my siblings, Mel and Biel; to my family and friends around the world. Thank you for the support and for understanding my lack of communication and visits. To my parents-in-law, Jaap and Jenny, for being my second family, comforting and helping me daily and making me feel at home. To my cats, Sybil and Manuel, thank you for breaking the silence during my long working days, reminding me that there is life outside.

To my husband Gijsbert and our twin girls, I owe you public apologies for not being around. Gijs, thank you for being a full time parent for our little ones when I was immersed in the dark side, too busy to share the parenthood duties. Thank you for holding our family together. Thank you for being the best husband, friend and English teacher in the galaxy. I would have never ever gone this far without you. You are my sunshine, you bring me to the bright side.

To Hannah and Nina, thank you for challenging me more than each chapter of this thesis. Raising you two makes writing a thesis seem like an easy task, as much as writing a thesis makes motherhood seem easy. Thank you for reminding me every day that dreams have no limits, but big dreams demand big journeys. I hope I have set an example of a believer, dreamer and hard worker for you both.

For everyone who is reading this: May the force be with you.

A

CURRICULUM VITAE

Camila Caram Deelder was born in 1977 in Belo Horizonte, Brazil. She was educated in a Montessori school where she could comfortably divide her attention between an unusual mixture of interests in maths, computers, "medicine" and handcrafts. In 2000 she graduated as Bachelor of Sciences in Statistics with emphasis on biostatistics at the Institute of Exact Sciences at Universidade Federal de Minas Gerais. In 2006 she received her Master of Science in Public Health, with emphasis on Epidemiology, at the Faculty of Medicine of the same university. On her master project, Camila mapped and analysed the space-temporal distribution of blood donors in two time points 10 years apart.

Since her bachelor graduation Camila has been working as consultant and teacher for institutes, universities and governmental agencies collaborating on projects on the fields of Epidemiology, Statistics and Data Management in Brazil, Canada, England, Portugal and the Netherlands. As a permanent contractor she worked for an Internet service provider (Planetarium, Brasil) in 1996 and 1997, for a research company (Vox Populi, Vox do Brasil) from 2001 to 2003 and for a health insurance company (Unimed-BH, Brasil) from 2006 to 2008.

In 2008 Camila moved to London to study English and expand her project collaborations in Europe. In 2009 she worked in a collaboration with Prof. Dr. Anske van der Bom and Dr. Suely Meireles Rezende on a project regarding Factor VIII Inhibitors in hemophilia A patients. In this occasion she visited the Department of Clinical Epidemiology of Leiden University Medical Center as a researcher.

In 2011 Camila was invited to work on a one year project on *sex-mismatched transfusions and recipient mortality* which was a collaboration between the Dutch Bloodbank (Sanquin) and the Leiden University Medical Center. This project was later expanded and was incorporated into this thesis.

During her thesis years Camila deepened her passion for databases and management of secondary data. She found in "Data Stewardship" the opportunity to transform data into knowledge. With that she found the career that she wanted to follow. In January 2017 she started working on her new job position as data stewardess at the same department where she is graduating. Since then Camila has already worked in several researches, PhD and Master projects performed by the Sanquin and the Leiden University Medical Center.

Camila is married to Gijsbert. They are parents of twins girls Hannah and Nina and live in Voorhout, the Netherlands, with their cats Sybil and Manuel.

PROPOSITIONS – STELLINGEN

THE BRIGHT AND THE DARK SIDE OF BLOOD TRANSFUSION
TURNING DATA INTO KNOWLEDGE

01. *Blood from female donors who have been pregnant increases mortality of male patients.* (this thesis)

02. *For the sake of science 'guidelines makers' and 'guidelines users' should reach a consensus about the terms 'prospective' and 'retrospective' in observational study reports.* (this thesis)

03. *Old platelets are inferior to fresh platelets for all post transfusion platelet measurements, transfusion interval and the need for additional platelet transfusions.* (this thesis)

04. *There is no evidence that storage time of platelets affects the number of transfusion reactions, bleeding and need for transfusions of any blood component.* (this thesis)

05. *The gain of giving only fresh platelets is unlikely to outweigh the benefit of extended storage time.* (this thesis)

06. Platelets transfusions can fail due to storage time only, irrespective of any patient characteristics or clinical factors, resulting in unnecessary diagnoses of refractoriness. (this thesis)

07. *If the true magnitude of bias is similar to the difference in our point estimates, then selection bias has the potential to change study conclusions.* (Power et. al., Epidemiology 2013)

08. *A transfusion should never be ordered or given, unless it is worth the risk.* (Karl Landsteiner, 1868-1943)

09. *Red blood cells were never intended to be put into a plastic bag and kept in a fridge for weeks; neither were platelets ever intended to be put into a plastic bag and kept at room temperature for days!* (paraphrasing Leo van de Watering, 2012)

10. *Absence of evidence is not evidence of absence.* (Altman & Bland, BMJ, 1995)

11. *At the start of every disaster movie there is a scientist being ignored.* (March for Science, 2017 – unfortunately the power of making decisions is not in the hands of scientists)

12. *A designer knows he has achieved perfection not when there is nothing left to add, but when there is nothing left to take away.* (Antoine de Saint-Exupery, 1939 – scientists are designers of mathematical models and make subjective decisions during the design process)

13. *The hottest places in hell are reserved for those who, in a period of moral crisis, maintain their neutrality.* (John F. Kennedy, 1958 – good scientists will always choose a side based on the evidence they have)

14. *Do or do not, there is no try.* (Master Yoda, 1980 – to achieve goals, your mind must be fully committed or you will not be able to do it)

Camila Caram Deelder
Leiden, 2017