International validation of harmonized definitions for complications of blood donations

Kevin J. Land,[1,2] *Mary Townsend*,[1] *Mindy Goldman*,[3] *Barbee I. Whitaker*,[4] *Gabriela E. Perez*,[4] *and Jo C. Wiersum-Osselton*[5]

BACKGROUND: In December 2014, a multinational collaboration of hemovigilance experts from the International Society of Blood Transfusion (ISBT), the International Hemovigilance Network, and AABB published harmonized definitions of complications related to blood donation titled "Standard for Surveillance of Complications Related to Blood Donation." Both mandatory and optional terms were included. The definitions are endorsed by the Alliance of Blood Operators and the European Blood Alliance.

STUDY DESIGN AND METHODS: The objective of this study was to validate harmonized donor hemovigilance definitions with potential users. In June 2016, 30 real-world cases were sent to potential users around the world along with the definitions, an answer sheet, and instructions on how to complete the validation exercise.

RESULTS: Overall, 54 responses from 25 countries were received, including over 400 comments. The results were presented for feedback at both ISBT and AABB meetings. Case diagnoses were consistent across most responders. Exceptions were rare adverse events, nonstandard presentations, or incomplete information. In general, the application of optional definitions, including severity grading and imputability, had the most variability.

CONCLUSION: The use of standardized terms in the donor setting serves to increase focus on donor safety, facilitate conversation, foster exchange of information, and frame questions for future research. Overall, the definitions provide adequate coverage of donor reactions; however, some terms require clarification. Severity grading and imputability and other optional terms need clear and objective definitions and instructions on when and how to use them. Additional feedback and final recommendations are summarized in this report.

The practice of hemovigilance (HV) began with assessing the frequency of undesired or unexpected effects of transfusion in recipients.[1-3] More recently, increased attention has focused on adverse events associated with blood donation.[4-6] Blood donation is extremely safe, although very rarely it is associated with serious donor injury or long-term disability. Less severe adverse events can be problematic, as they may yield an insufficient collection volume or an unpleasant donation experience, potentially reducing donor return. Blood collection agencies use various mitigation strategies to reduce adverse events, including deferring donors at highest risk for adverse events, encouraging predonation fluid and salt loading, using applied muscle tension, optimizing the collection staff and environment, and ameliorating donation-associated fear and anxiety.[7-14] Beyond establishing baseline reaction rates (e.g., benchmarking), a robust donor HV system is necessary to assess the impact of diverse operational changes and mitigation strategies and to facilitate the reliable exchange of information and ideas.

ABBREVIATIONS: HV = hemovigilance; IHN = International Haemovigilance Network; ISBT = International Society of Blood Transfusion

From [1]Blood Systems, Inc., and [2]UT Health Science Center at San Antonio, Department of Pathology, San Antonio, Texas; [3]Canadian Blood Services, Ottawa, Ontario, Canada; [4]AABB, Bethesda, Maryland; and [5]TRIP Hemovigilance and Biovigilance, Leiden, the Netherlands.

Address reprint requests to: Kevin J. Land, MD, Adjunct Professor of Pathology, UT Health Science Center at San Antonio Vice President of Clinical Services, Blood Systems, Inc., Medical Director, Cellular Therapy, Blood Systems Laboratory, 6210 East Oak Street, Scottsdale, AZ 85257-1101; e-mail: kland@bloodsystems.org

Received for publication May 25, 2018; revision received July 26, 2018; and accepted July 26, 2018.

doi:10.1111/trf.14948

TRANSFUSION 2018;58;2589–2595

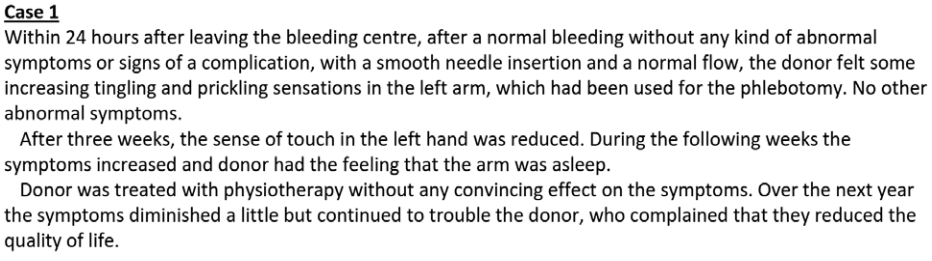
The cornerstone of any vigilance system is a clear, uniformly applied set of definitions. In 2014, the AABB Donor Hemovigilance Working Group, the International Haemovigilance Network (IHN), and the International Society of Blood Transfusion (ISBT) Haemovigilance Working Party formed a Consensus Working Group that merged their existing HV definitions into a single, updated, harmonized donor reaction definitions document.[15-17] Because HV systems vary in level of detail, in addition to basic mandatory categories, the definitions included optional elements, allowing for subdivision of categories or additional information such as the site of the reaction. Once the definitions were endorsed by AABB, the IHN, ISBT, and other organizations, 30 short adverse event case studies were employed for use in a global validation exercise to assess concordance of the use of these definitions among current donor HV users. The responses were initially discussed at the ISBT Haemovigilance Working Party meeting in Dubai, United Arab Emirates, in September 2016. Feedback from this discussion was summarized and shared with active HV members from ISBT, IHN, and AABB for additional comments and recommendations. The outcome of the validation exercise is reported in this article.

MATERIALS AND METHODS

The objective of this study was to validate recently harmonized donor HV definitions with potential users using real-world cases collected from active international HV members of ISBT, IHN, and AABB. A few cases associated with rare adverse events were constructed from the published case literature by one author. Thirty short cases were selected to broadly represent the spectrum of adverse events, their presentations, associated signs and symptoms, and severity. All categories were addressed by the cases except for the following rare diagnoses: apheresis-related hemolysis (C2), apheresis-related air embolism (C3), myocardial infarction (E2), cardiac arrest (E3), and death (E6). To evaluate the real-world applicability of the definitions, cases were only minimally edited to provide variety in either the case presentation or follow-up. Each case consisted of a brief description of the donor and procedure, a description of the adverse event, and any follow-up. Most of the cases (20 of 30) contained a single adverse event. Case 1 is provided as an example (Fig. 1).

In June 2016, the cases and associated documents, including a cover sheet requesting universal participation, instructions, examples of cases with potential answers, an answer scheme, a formatted answer sheet, and the harmonized definitions, were sent out to active members of the AABB and ISBT HV groups. It was also distributed more generally via e-mail with support from ISBT, IHN, AABB, and America's Blood Centers. The participants were asked in the cover letter to evaluate each case, selecting at least one adverse donor category for each case and evaluating the optional categories or attributes of that diagnosis where appropriate. For each answer, at least one diagnosis was requested along with any associated optional information, such as duration of arm pain, length of loss of consciousness, severity and imputability scores, and any comments regarding the use of the harmonized definitions. The answer scheme showing mandatory categories and optional categories (with asterisks and in italics), as well as the severity grading and imputability scale (likelihood that adverse event was caused by donation) is presented in Fig. 2. All documents associated with the validation exercise, including a PowerPoint file summarizing the responses for each case and other resource files are available on the ISBT and AABB Web sites.[16,17]

Descriptive statistics were obtained to identify the most frequent definition, severity, and imputability category in each donor case. Interrater reliability was calculated using Fleiss kappa (κ) statistics to measure the level of agreement or concordance among all responders and the level of agreement by donor adverse event (category-wise κ). Donor cases were grouped by reaction categories A through F (Table 1). The level of agreement (κ) was divided into quintiles for interpretation: Poor ($\kappa \leq 0.20$), Fair ($0.20 < \kappa \leq 0.40$), Moderate ($0.40 < \kappa \leq 0.60$), Good ($0.60 < \kappa \leq 0.80$), and Very Good ($\kappa > 0.80$). Only one category (A: Local Symptoms) was used to calculate κ in cases with multiple potential diagnoses. Interrater statistics were not calculated when there were not at least two raters and at least two cases. Statistical analysis was performed using computer software R (R Foundation for Statistical Computing, Rv3.3.1), and kappa statistics with a threshold of p less than 0.05 were considered significant.



Case 1

Within 24 hours after leaving the bleeding centre, after a normal bleeding without any kind of abnormal symptoms or signs of a complication, with a smooth needle insertion and a normal flow, the donor felt some increasing tingling and prickling sensations in the left arm, which had been used for the phlebotomy. No other abnormal symptoms.

After three weeks, the sense of touch in the left hand was reduced. During the following weeks the symptoms increased and donor had the feeling that the arm was asleep.

Donor was treated with physiotherapy without any convincing effect on the symptoms. Over the next year the symptoms diminished a little but continued to trouble the donor, who complained that they reduced the quality of life.

Fig. 1. Sample case for review.

Fig. 2. 2016 AABB, International Haemovigilance Network, and ISBT donor hemovigilance case validation answer scheme (optional categories with asterisks and in italics).

RESULTS

A total of 54 responses were received from 25 countries. Based on self-reporting demographics, the respondees worked predominantly in either academic institutions, governmental agencies, or blood centers and mostly had titles that included words such as Professor, Medical Director, Deputy, or Director (data not shown). Four respondees stated that their response reflected input from staff or fellow or junior faculty. At least one severity and imputability result was provided by 52 of 54 responders for each case. The interrater reliability was calculated for 18 of the unique diagnosis cases and for nine of the multiple-diagnosis cases.

Cases with a single diagnosis

The level of agreement between responders is shown in Table 1. Based on Fleiss κ analysis, Category A: Local Symptoms cases had overall good agreement ($\kappa = 0.742$; $p < 0.001$). Arterial puncture (A1.2), cellulitis (A3.2), deep venous thrombosis (A4.1), and compartment syndrome (A4.3) showed very good agreement (>0.80 κ; p

Cases with multiple diagnoses

When further analysis was done for multiple-diagnosis cases that included Local Symptoms, an overall moderate agreement was observed ($\kappa = 0.449$). Delayed bleeding (A1.3) and arteriovenous fistula (A4.2) had good agreement levels (0.608 and 0.772 category-wise κ, respectively).

Optional responses and comments

In one-half of the cases (15 of 30), at least one optional subcategory or term was possible in the answer sheet (Fig. 2), excluding severity and imputability (discussed below). Use of optional subcategories varied. Optional categories, such as superficial thrombophlebitis (A3.1) and cellulitis (A3.2), were used by nearly all responders (53 of 54 and 52 of 54, respectively), whereas responders less often made use of other arm pain (A2.2). Responder comments suggested that the other arm pain category was confusing to them (data not shown).

Optional descriptors included duration of symptoms, presence/absence of injury, and onsite/offsite location of reaction. In general, the descriptors were more likely to be used in cases where the descriptor was explicitly mentioned and if only one optional descriptor was requested. For example, in cases diagnosed as nerve injury/irritation (A2.1), 76% to 94% of responders reported the single optional descriptor duration of symptom. For vasovagal reactions, however, there was more variability. Duration of loss of consciousness was most often reported (range 54%-87%) followed by presence of injury (range, 52%-72%) and location of reaction (61%-70%). Injury was more likely to be

TABLE 1. Single-diagnosis cases: level of agreement between assessors

Diagnosis				Category-wise kappa	Level of agreement
A. Complications mainly with local symptoms	Blood outside vessel	A1.1	Hematoma	0.791	Good
		A1.2	Arterial puncture	0.847	Very good
		A1.3	Delayed bleeding	NS*	
	Arm pain	A2.1	Nerve injury/irritation	0.787	Good
		A2.2	*Other arm pain*†	*0.19*	*Poor*
	Localized infection/ inflammation	*A3.1*	*Superficial thrombophlebitis*†	*0.716*	*Good*
		A3.2	*Cellulitis*†	*0.810*	*Very good*
	Other major blood vessel injury	A4.1	Deep venous thrombosis	0.833	Very good
		A4.2	Arteriovenous fistula	0.101	Poor
		A4.3	Compartment syndrome	0.833	Very good
		A4.4	Brachial artery pseudoaneurysm	0.712	Good
B. Generalized symptoms—vasovagal reactions		B1	Vasovagal reaction, no loss of consciousness	0.728	Good
		B2	Vasovagal reaction, loss of consciousness	0.743	Good
C. Related to apheresis		C1	Citrate reactions	0.796	Good
		C2	Hemolysis	NS*	
		C3	Air embolism	NS*	
		C4	Infiltration	0.512	Moderate
D. Allergic reactions		D1	Local allergic reaction	NS*	
		D2	Generalized reaction	NS*	
E. Other serious conditions		E1	Acute cardiac symptoms (other than E2 or E3)	0.491	Moderate
		E2	Myocardial infarction	NS*	
		E3	Cardiac arrest	NS*	
		E4	Transient ischemic attack	0.256	Fair
		E5	Cerebrovascular accident	0.304	Fair
		E6	Death	NS*	
F. Other		F	Other	NS*	

* Represents categories with nonsignificant kappa.
† Italicized terms are optional.

reported than "without injury." Similarly, location of reaction (onsite vs. offsite) was more likely to be reported if explicitly stated in the scenario (data not shown).

There were approximately 400 total free text comments submitted, averaging seven comments per responder (range, 0-30). Each question averaged 13 free text comments (range, 8-26). In general, the number of free text comments correlated with the complexity of the case (data not shown).

Severity coding varied widely. Figure 3 demonstrates the range of assigned severity in a group of cases where arm pain is the best single diagnosis. These range from an assignment of Mild by 93% of responders for Case 22 to varying assignments of severe (46%), moderate (44%), and mild (6%) for Case 8. Figure 4 demonstrates similarly variable responses received for imputability among the three cases with vasovagal reaction as a single diagnosis, where for Case 16, 7% of responders thought the case was "definitely" related to the donation compared with 72% of responders who found the case unlikely to be or excluded from being related to the donation event. Case 21 results showed similar variation, with 22% "probably" related to the donation, and another 20% found it unlikely to have been related to the donation.

DISCUSSION

In this validation study, users from around the world were asked to evaluate a harmonized donor HV definition list that was created from existing lists by a multinational collaboration of HV experts. Overall, the validation exercise successfully demonstrates that there is good agreement in how responders use the terms; however, there are several opportunities to clarify the definitions and when to use them.

A donor HV "ontology"

Health care is full of complicated terms and concepts that need to be defined precisely and objectively. Without standardized, well-defined terminology, communication is poor, reducing our ability to learn from others and translate that new knowledge into local continuous process improvement initiatives. National and global surveillance systems are important ways to standardize terminology and define commonly reported data elements but struggle with the opposing goals of promoting comprehensive versus providing simple reporting.

Comprehensive reporting helps gather a wide variety of detailed information that is necessary to exchange complicated ideas and can lead to hypothesis generation or better understanding of underlying mechanisms. Unfortunately, it

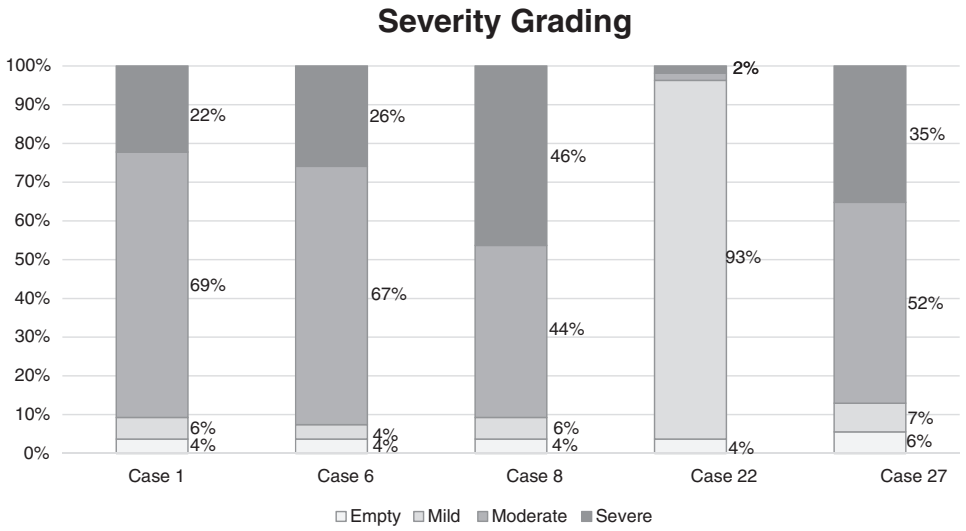


Fig. 3. Range of severity rating in cases of arm pain (54 responses per case).

is very resource intense. With simple reporting requirements, on the other hand, it is less burdensome to collect information, and the information is more likely to be captured electronically at the expense of the collecting much less information. Policy makers increasingly need real-time data to inform policy, requiring the use of clearly and objectively defined terms across large systems and geographic regions that are electronically captured for quick retrievability.[18] Too often, however, the real questions begin where the simple data reporting ends.

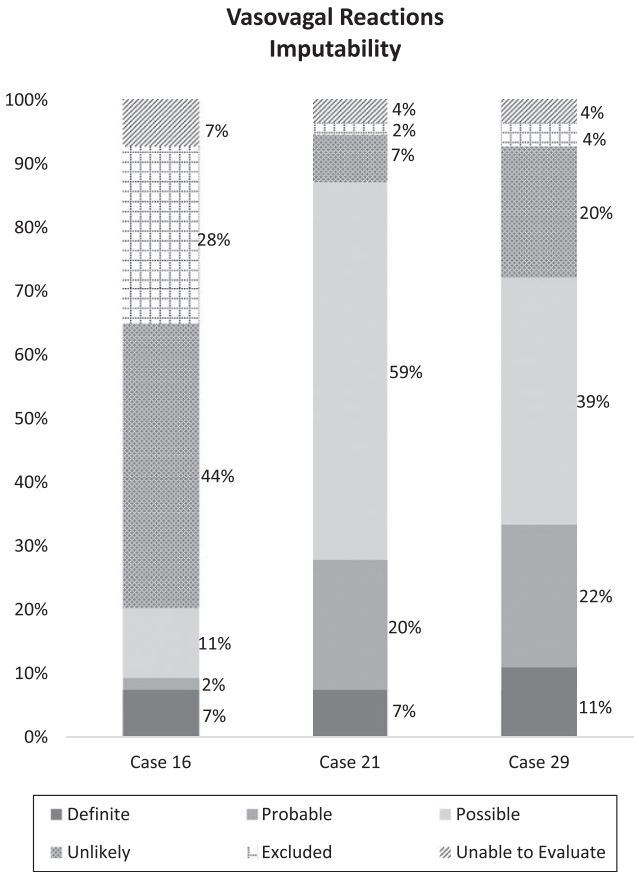


Fig. 4. Range of imputability rating in cases of vasovagal reaction (54 responses per case).

Ontologies provide a unifying framework that is conducive to problem solving and are defined by the *Oxford English Dictionary* as "a set of concepts and categories in a subject area or domain that shows their properties and relations between them."[19,20] The current harmonized definitions list is loosely organized as an ontology, providing both the ability to capture detailed information in a structured way for local or research use that can be condensed into at least two more generalized levels (Fig. 5) better suited for surveillance systems. In the validation exercise, the agreement of the definitions generally improved as the terms were integrated into more general categories, suggesting that future revisions should provide more guidance to the definitions and uses of the granular terms. The validation exercise also shows us that HV cases must contain sufficient detail and objective information to be accurately reported.

Optional versus mandatory terms

The addition of optional terms in the ontology serves many purposes: (1) It was considered the best way to organize information that may have utility at a local or academic level that could readily be mapped to broader surveillance terms; (2) it allowed flexibility in incorporating terms that may be used more frequently in some regions over others (such as the more common usage of the term *infiltration* [C4] in North America); and (3) it provided a framework to introduce potentially emerging concepts into the ontology. Other arm pain (A2.2), for example, is beginning to be used in certain western European countries to try to distinguish

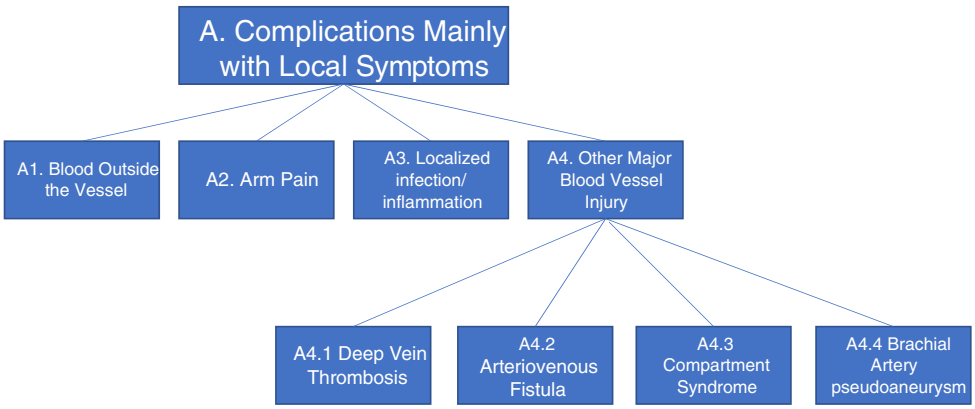


Fig. 5. Ontological Structure Sample: Complications Mainly with Local Symptoms. [Color figure can be viewed at wileyonlinelibrary.com]

direct nerve damage with vague pain that quickly resolves. Additionally, we need an opportunity to introduce new terms as we continue to learn about adverse donor events. The ontology should remain flexible to include or at least consider additional data elements that might be helpful in improving the objectivity of terms.

Multiple diagnoses

Future revisions will need to address how to manage reporting of two or more separate but related complications, for instance, when a hematoma results in arm pain. Phlebotomy involves a sharp needle piercing both soft tissue and veins, and cutaneous nerve fibers and arterial vessels run alongside veins (and sometimes around); therefore, it is not surprising that blood outside vessels (A1) is often associated with arm pain (A2), vasovagal reactions (B1, B2), and other donor-related adverse events. The question becomes, "Is it important to report one complication over the others, or to capture all of them?" From the donor's perspective, it is likely considered worse to have multiple adverse events, whereas HV systems tend to focus on the complication of greater severity or longest duration. The answer, therefore, may depend on how the information is to be used and should require objective definition and clear distinction between levels of severity and imputability.

Capturing the evolution of an adverse event

Symptoms may evolve over the first 24 hours, especially when associated with rebleeding or delayed bleeding, and so are not always fully captured at presentation. Not all blood establishment computer systems are capable of capturing follow-up information, meaning that even if there is an automated way to capture initial data, a significant portion of donor HV data must still be manually captured in a separate database. For these and other reasons, at least in the United States, adoption of automated data capture of donor HV has been slow.[6]

Severity and imputability

Assigning severity and imputability can be difficult, especially when information is incomplete, and some terms, such as *long-term pain and/or disability,* are subjective. Both severity and imputability assessments were intentionally made optional during the initial version, as there are no uniformly agreed upon objective criteria to separate levels of severity or imputability. This flexibility allows blood centers to determine which elements of donor vigilance are feasible in their situation; however, it limits the ability to combine data into larger data sets that would likely better inform policy.

Current definitions of severity are more aligned to recipient reactions,[21,22] and may be difficult to directly apply to donor complications. Yet, a consensus may be gathering that there are some objective criteria—such as the need for outside medical care, the need for immediate surgery, and temporary or permanent changes in level of donor function—that are documented consistently enough to be considered in future revisions.

TABLE 2. General suggestions for subsequent revisions

1. Utilize a format such as the recipient adverse event template to more clearly structure the terms.
2. Incorporate suggestions to clarify terms (e.g., change *deep venous thrombosis* to *venous thrombosis*).
3. Include examples of when to use and not use specific terms, including the optional terms.
4. Define terms and levels of severity and imputability in reproducibly objective formats.
5. Determine when and how multiple diagnoses should be reported with examples.
6. Consider if other terms (e.g., needle adjustment) or dictionaries (e.g., Plasma Protein Therapeutics Association definitions) should be incorporated.

Next steps

Based on comments from responders and feedback during presentations, the first version of the harmonized definition list has successfully passed its validation exercise. It is almost universally agreed that a single list is preferred when sharing donor HV data internationally and that future revisions should focus on setting objective criteria for the assignment of severity and imputability. The authors have collected many suggestions during this exercise and have included them in the resource documents available on the AABB and ISBT Web sites. The topics are broadly summarized in Table 2.

In conclusion, the current harmonized donor HV terms adequately cover donor HV diagnoses. The overall concordance in use of the terms is good and in general improves when there is relative clarity on symptoms to be reported and instructions for use of the terms.

ACKNOWLEDGMENTS

The authors thank all participants for their input, as well as everyone who has provided feedback and suggestions since the validation exercise began. Author contributions are as follows: KJL, first author; MT, cases; GEP, statistical analysis; MG, secondary author; and JCWO, study design and manuscript review.

CONFLICT OF INTEREST

The authors have disclosed no conflicts of interest.

REFERENCES

1. Land KJ, Whitaker BI, Uhl L. National hemovigilance: the current state. AABB Technical Manual. 19th ed. Bethesda (MD): AABB; 2018.
2. de Vries RR, Faber JC, Strengers PF. Haemovigilance: an effective tool for improving transfusion practice. Vox Sang 2011;100:60-7.
3. AuBuchon JP, Fung MK, Perez G, et al. Classification of post-transfusion adverse events using a web-based algorithm. Transfusion 2017;57:489-90.
4. Wiersum-Osselton JC, Marijt-van der Kreek T, de Kort W. Donor vigilance: what are we doing about it? Biologicals 2012;40:176-9.
5. Sorensen BS, Johnsen SP, Jorgensen J. Complications related to blood donation: a population-based study. Vox Sang 2007;94:132-7.
6. Rajbhandary S, Stubbs JR, Land KJ, et al., editors. AABB donor hemovigilance report: 2012–2014 on behalf of the AABB US Donor Hemovigilance Working Group. 2016 [cited 2018 Apr 4]. Available from: http://www.aabb.org/research/hemovigilance/documents/2012-2014-AABB-Donor-Hemovigilance-Report.pdf.
7. Sachjev S, Sing L, Sharma RR, et al. A study on the effect of pre-donation salt loading on vasovagal reactions in young college going whole blood donors. Indian J Hematolog Blood Transfus 2017;33:592-7.
8. France CR, Ditto B, Wissel ME, et al. Predonation hydration and applied muscle tension combine to reduce presyncopal reactions to blood donation. Transfusion 2010;50:1257-64.
9. Hanson SA, France CR. Social support attenuates presyncopal reactions to blood donation. Transfusion 2009;49:843-50.
10. Eder AF, Dy BA, Kennedy JM, et al. Improved safety for younger whole blood donors with new selection criteria for total estimated blood volume. Transfusion 2011;51:1522-31.
11. Morand C, Coudurier N, Rolland C, et al. Prevention of syncopal-type reactions after whole blood donation: a cluster-randomized trial assessing hydration and muscle tension exercise. Transfusion 2016;56:2412-21.
12. Fisher SA, Allen D, Dorée C, et al. Interventions to reduce vasovagal reactions in blood donors: a systematic review and met-analysis. Transfusion Med 2016;26:15-33.
13. Goldman M, Osmond L, Yi QL, et al. Frequency and risk factors for donor reactions in an anonymous blood donor survey. Transfusion 2013;53:1979-84.
14. Bravo M, Kamel H, Custer B, et al. Factors associated with fainting—before, during and after whole blood donation. Vox Sang 2011;101:303-12.
15. Goldman M, Land K, Robillard P, et al. Development of standard definitions for surveillance of complications related to blood donation. Vox Sang 2016;110:185-8.
16. ISBT website [cited 2018 Mar 16]. Available from http://www.isbtweb.org/working-parties/haemovigilance/(click on definitions).
17. Working Group on Donor Vigilance of the International Society of Blood Transfusion Working Party on Haemovigilance, The International Haemovigilance Network, The AABB Donor Haemovigilance Working Group. Standard for surveillance of complications related to blood donation. December 11, 2014 [cited 2018 Apr 4]. Available from: http://www.aabb.org/research/hemovigilance/Documents/Donor-Standard-Definitions.pdf.
18. Public Health Service Biovigilance Task Group Biovigilance in the United States: efforts to bridge a critical gap in patient safety and donor health. Washington (DC): HHS; 2009 [cited 2018 Apr 6]. Available from: https://wayback.archive-it.org/3922/20140403203201/http://www.hhs.gov/ash/bloodsafety/biovigilance/ash_to_acbsa_oct_2009.pdf.
19. Uschold M, Gruninger M. Ontologies: principles, methods, and applications (AIAI-TR-191). Knowl Eng Rev 1996;11:1-69.
20. OED Online. January 2018 [cited 2018 Apr 6]. Oxford University Press. Available from: https://en.oxforddictionaries.com/definition/ontology.
21. Wiersum-Osselton JC, Faber JC, Politis C, et al. Quality validation of data in national haemovigilance systems in Europe: report of a survey on current state of practice. Vox Sang 2013;104:214-7.
22. AuBuchon JP, Fung M, Whitaker B, et al. AABB validation study of the CDC's National Healthcare Safety Network Module adverse events definitions protocol. Transfusion 2014;54:2077-83.