Available online at www.annclinlabsci.org
Annals of Clinical & Laboratory Science, vol. 45, no. 2, 2015

Post-Operative Bleeding Risk Stratification in Cardiac Pulmonary Bypass Patients Using Artificial Neural Network

Richard S.P. Huang[1], Elena Nedelcu[1], Yu Bai[1], Amer Wahed[1], Kimberly Klein[1], Hlaing Tint[1], Igor Gregoric[2], Manish Patel[2], Biswajit Kar[2], Pranav Loyalka[2], Sriram Nathan[2], Rajko Radovancevic[2], and Andy N.D. Nguyen[1]

[1]The University of Texas Health Science Center at Houston-Department of Pathology and Laboratory Medicine and [2]The University of Texas Health Science Center at Houston-Center for Advanced Heart Failure, Houston, TX, USA

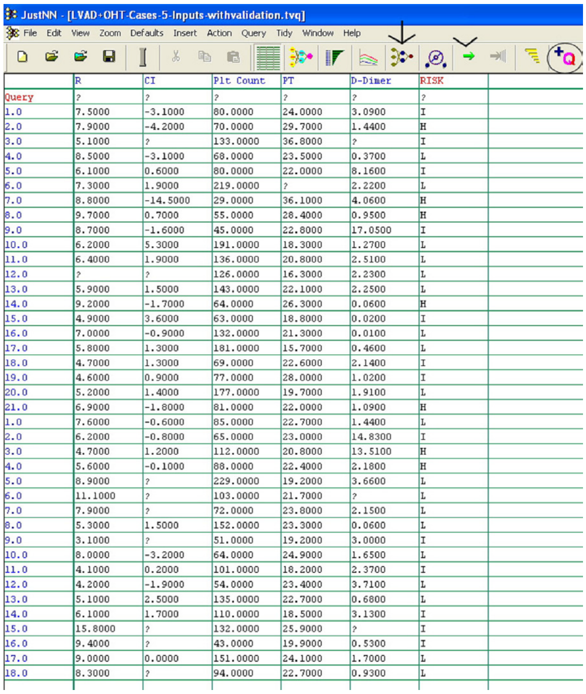
Abstract. The prediction of bleeding risk in cardiopulmonary bypass (CPB) patients plays a vital role in their postoperative management. Therefore, an artificial neural network (ANN) to analyze intra-operative laboratory data to predict postoperative bleeding was set up. The JustNN software (Neural Planner Software, Cheshire, England) was used. This ANN was trained using 15 intra-operative laboratory parameters paired with one output category – risk of bleeding, defined as units of blood components transfused in 48 hours. The ANN was trained with the first 39 CPB cases. The set of input parameters for this ANN was also determined, and the ANN was validated with the next 13 cases. The set of input parameters include five components: pro-thrombin time, platelet count, thromboelastograph-reaction time, D-Dimer, and thromboelastograph-coagulation index. The validation results show 9 cases (69.2%) with exact match, 3 cases (23.1%) with one-grading difference, and 1 case (7.7%) with two-grading difference between actual blood usage versus predicted blood usage. To the best of our knowledge, ours is the first ANN developed for post-operative bleeding risk stratification of CPB patients. With promising results, we have started using this ANN to risk-stratify our CPB patients, and it has assisted us in predicting post-operative bleeding risk.

Introduction

Our pathology department started a new service called hemotherapy in May 2012. The hemotherapy team (HT), which is composed of hematopathologists, blood bank pathologists, and pathology residents, provides consultation to cardiac surgeons and anesthesiologists on coagulation issues intra-operatively and post-operatively. Their patients undergo cardiopulmonary bypass (CPB) surgery for left ventricular assisted device (LVAD) and cardiac transplant surgery. The recommendations made by the HT to the anesthesiologist and cardiac surgeons are based on laboratory data that are run in the STAT laboratory next to the operating room. Intraoperatively, the HT formulates a recommendation of blood products to transfuse for the anesthesiologist and cardiac surgeons based on this laboratory data and the HT's clinical impression of the patient. Postoperatively, the HT takes charge of the management of the patient's bleeding by ordering blood products to help control the bleeding [1].

During post-operative care, the HT finds that prediction of bleeding risk plays a vital role in patients' management. Therefore, members of the HT sought to develop a computerized program to predict postoperative bleeding. With some previous experience in using artificial neural networks (ANN), the HT experimented with this idea. ANN are mathematical/computational programs that are modeled on the biological central nervous system (neural networks) and are composed of interconnecting nodes (neurons) that can recognize patterns and relationships in data. An ANN can take input data/output data and undergo supervised learning with a back-propagation method, or it can take input data only and undergo unsupervised learning with an adaptation or self-organization method [2]. While ANN is used as a tool in a variety of settings outside of medicine, in the medical setting, it has been primarily used for classification, pattern recognition, prediction, and modeling [2].

Address correspondence to Richard S.P. Huang, MD and Andy N.D. Nguyen, MD; The University of Texas Health Science Center at Houston, Department of Pathology and Laboratory Medicine, 6431 Fannin Street MSB 2.292, Houston, Texas 77030; phone: (713) 500 5337; fax: (713) 500 0712; e mails: Richard.Huang.1@uth.tmc.edu and Nghia.D.Nguyen@uth.tmc.edu



	R	CI	Plt Count	PT	D-Dimer	RISK
Query	?	?	?	?	?	?
1.0	7.5000	-3.1000	80.0000	24.0000	3.0900	I
2.0	7.9000	-4.2000	70.0000	29.7000	1.4400	H
3.0	5.1000	?	133.0000	36.8000	?	I
4.0	8.5000	-3.1000	68.0000	23.5000	0.3700	L
5.0	6.1000	0.6000	80.0000	22.0000	8.1600	I
6.0	7.3000	1.9000	219.0000	?	2.2200	L
7.0	8.8000	-14.5000	29.0000	36.1000	4.0600	H
8.0	9.7000	0.7000	55.0000	28.4000	0.9500	H
9.0	8.7000	-1.6000	45.0000	22.8000	17.0500	I
10.0	6.2000	5.3000	191.0000	18.3000	1.2700	L
11.0	6.4000	1.9000	136.0000	20.8000	2.5100	L
12.0	?	?	126.0000	16.3000	2.2300	L
13.0	5.9000	1.5000	143.0000	22.1000	2.2500	L
14.0	9.2000	-1.7000	64.0000	26.3000	0.0600	H
15.0	4.9000	3.6000	63.0000	18.8000	0.0200	I
16.0	7.0000	-0.9000	132.0000	21.3000	0.0100	L
17.0	5.8000	1.3000	181.0000	15.7000	0.4600	L
18.0	4.7000	1.3000	69.0000	22.6000	2.1400	I
19.0	4.6000	0.9000	77.0000	28.0000	1.0200	I
20.0	5.2000	1.4000	177.0000	19.7000	1.9100	L
21.0	6.9000	-1.8000	81.0000	22.0000	1.0900	H
1.0	7.6000	-0.6000	85.0000	22.7000	1.4400	L
2.0	6.2000	-0.8000	65.0000	23.0000	14.8300	I
3.0	4.7000	1.2000	112.0000	20.8000	13.5100	H
4.0	5.6000	-0.1000	88.0000	22.4000	2.1800	H
5.0	8.9000	?	229.0000	19.2000	3.6600	L
6.0	11.1000	?	103.0000	21.7000	?	L
7.0	7.9000	?	72.0000	23.8000	2.1500	L
8.0	5.3000	1.5000	152.0000	23.3000	0.0600	L
9.0	3.1000	?	51.0000	19.2000	3.0000	I
10.0	8.0000	-3.2000	64.0000	24.9000	1.6500	L
11.0	4.1000	0.2000	101.0000	18.2000	2.3700	I
12.0	4.2000	-1.9000	54.0000	23.4000	3.7100	L
13.0	5.1000	2.5000	135.0000	22.7000	0.6800	L
14.0	6.1000	1.7000	110.0000	18.5000	3.1300	I
15.0	15.8000	?	132.0000	25.9000	?	I
16.0	9.4000	?	43.0000	19.9000	0.5300	I
17.0	9.0000	0.0000	151.0000	24.1000	1.7000	L
18.0	8.3000	?	94.0000	22.7000	0.9300	L

Figure 1. *Grid view of 39 patient's laboratory values and bleeding risk in the Artificial Neural Network. Column 1: # 1-21 is the left ventricular assisted device patient numbers and #1-18 is the cardiac transplant patient numbers, listed consecutively from when the surgery was performed. Row 1: R (Reaction time in thromboelastography), CI (coagulation index in thromboelastography), Plt count (platelet count), PT (prothrombin time), RISK (RISK is defined as L [low, ≤13.5 units of blood component], I [intermediate, 13.5-22.5 units of blood component], and H [high, ≥22.5 units]). Arrow, grow new network icon; arrowhead, start learning icon; circle, query icon.*

Using laboratory data from CPB patients on the hemotherapy service stored on our mobile computing platform, members of the HT experimented with building an ANN with this data [3]. This article will focus on how the ANN was built.

Material and Methods

Patient Selection. This study was approved by the Committee for the Protection of Human Subjects (CPHS) of the University of Texas Health Science Center at Houston. The laboratory data from the first 52 consecutive CPB patients on the hemotherapy service (between May 2012 and October 2013) were examined for inclusion in this project. Specifically, the first 31 consecutive CPB patients were examined for inclusion in the NeuroXL Clusterization portion of the project. After the exclusion of 14 patients due to incomplete laboratory data, the remaining 17 patients were included for this portion of the project. Next, of the first 52 consecutive CPB patients on the hemotherapy service, the first 39 consecutive patients were used to create the ANN and the following 13 consecutive patients were used to validate the ANN.

Determining Cut-off points with NeuroXL Clusterizer. Before the supervised ANN was created, low, intermediate, and high blood usage cut-off points had to be determined. NeuroXL Clusterizer (OLSOFT LLC, Moscow, Russia), which is a Microsoft-Excel (Microsoft Corporation, Redmond, Washington) add-on program that uses an unsupervised neural network approach (Kohonen) for clustering data, was chosen for this task. The 15 laboratory parameters (laboratory values most indicative of a bleeding tendency among several sets of intraoperative test results was chosen) and a blood usage parameter (defined as units of blood products used in 48 hours from the start of surgery) for each of the first 31 consecutive CPB patients on our hemotherapy service were entered into an Excel worksheet. For this project, one unit of blood product was defined as one unit of packed red blood cell, one apheresis unit of platelets, one unit of fresh frozen plasma, or one dose of cryoprecipitate (pooled from 10 single units). Since NeuroXL Clusterizer requires complete data input for all training cases, all the cases in the first 31 consecutive CPB patients that lacked all 15 laboratory parameters were eliminated, leaving 17 cases to determine cut-off points in NeuroXL Clusterizer. Next, the previously installed NeuroXL Clusterizer was launched from the add-ins tab of MS-Excel, and the option to cluster the input data (17 cases with all 15 laboratory parameters) into 3 groups (low, intermediate, and high blood usage) was selected.

Building ANN with JustNN. Once the cut-off points for low, intermediate, and high blood usage were determined, the process to build the ANN began. For this, we used JustNN software (Neural Planner Software, Cheshire, England), a program that can be downloaded free of charge from http://www.justnn.com/ and installed on any Windows-based computer to build our ANN. This program performs supervised training with input and output data points by using a back-propagation technique. When starting the JustNN program, it displays a screen similar to an Excel spreadsheet **(Figure 1)** where an ANN can be created by entering in all data and creating labels for rows and columns. In this particular ANN, each row is labeled with a case number and is followed by the laboratory data for that particular case. Each column is labeled with a particular laboratory parameter and is followed by the particular laboratory parameter for each individual case. The last column is labeled as the risk of bleeding or blood usage (output category) and is followed by the risk of bleeding or blood usage for each individual patient. The last column (blood usage) was designated as the output category by double-clicking any cell in the column and changing it to "output", thereby designating the whole column as an output column. Other columns associated with laboratory parameters are designated as "input". After all rows and columns were labeled and all laboratory data and output data from training cases were entered, a new network was created by clicking on the "grow new network" icon **(Figure 1)**. The network was then trained by clicking on the "start learning" icon **(Figure 1)**. Once the network had been created and the ANN trained, the "add query" icon was clicked **(Figure 1)** and a row labeled "query" was populated. The laboratory data of the patient that we wanted to query could then be entered into the query row. The computer program will then give a prediction of blood usage (the risk of bleeding) based upon the trained ANN in the row labeled "query" under the column labeled "Risk". Further details of how to create an ANN using JustNN can be found in the JustNN user guide [4].

Initially, the ANN was built and trained using 15 intraoperative laboratory parameters (thromboelastography-reaction time [TEG-R], heparinase modified thromboelastography-reaction time [hTEG-R], heparinase modified thromboelastography-Alpha [hTEG-A], heparinase modified thromboelastography-maximum amplitude [hTEG-MA], heparinase modified thromboelastography-estimated percent lysis [hTEG-EPL], heparinase modified thromboelastography-lysis at 30 minutes [hTEG-Ly30], heparinase modified thromboelastography-coagulation index [hCI], hemoglobin [Hgb], platelet [Plt], prothrombin time [PT], partial thromboplastin time [PTT], fibrinogen [Fib], thrombin time [TT], D-Dimer, and antithrombin III [ATIII]) paired with one output category, risk of bleeding, i.e. blood usage). Many sets of laboratory data are obtained during surgery (baseline, hemoconcentration phase/ warming phase, off-pump phase, and possible bleeding phase). The exact tests ordered in each set of laboratory tests and the number and timing of sets ordered are based on the clinical impression of the HT. However, only one laboratory value per parameter was chosen for use in building this ANN. Here, the parameter that is chosen is the most indicative of a bleeding tendency. For example, the highest PT/PTT value and the lowest platelet count among the result sets were chosen. For this project, the risk of bleeding is defined as units of blood components (one unit of packed red blood cell, one apheresis unit of platelets, one unit of fresh frozen plasma, or one dose of cryoprecipitate) transfused in 48 hours from the start of surgery: low, intermediate, and high based on the results of clustering by NeuroXL Clusterizer as described earlier.

Table 1. *Total Relative Importance of Input Parameters.*

	Parameter	Total Relative Importance
1	Platelets	303
2	Thromboelastograph-reaction time	244
3	Prothrombin time	204
4	D-Dimer	195
5	Thromboelastograph-coagulation index	169
6	Antithrombin III	133
7	VerifyNow-Plavix	115
8	Thromboelastograph-lysis at 30 minutes	77
9	Thromboelastograph- alpha	69
10	Fibrinogen	54
10	VerifyNow-Aspirin	54
11	Hemoglobin	32
12	Thromboelastograph-maximum amplitude	16
13	Partial thromboplastin time	9
14	Thrombin time	5

Table 2. *Validation Results*

Accuracy of Prediction	%
Exact match	69.2
One-grading difference	23.1
Two-grading difference	7.7
Accumulated exact match plus one-grading difference	92.3

The ANN was trained with the first 39 consecutive CPB cases (21 left ventricular assist device implants, and 18 heart transplants) on our hemotherapy service. The training was done with the default software settings, including learning rate of 0.60, momentum of 0.80, target error 0.01, and default setting for number of nodes [4]. The patient's input parameters would be entered in the query data entry field and the ANN would give a prediction on whether the patient was expected to need a low, intermediate, or high number of blood components. The prediction was then compared to the actual usage of blood components in the case.

The decision was made to perform subsetting of the data to reduce the input dimensions (number of input parameters). Of note, the ANN automatically calculates the relative importance of each input parameter. This is defined as the sum of the absolute weights of the connections from the input node to all the nodes in the first hidden layer. The subsetting process was performed as follows:

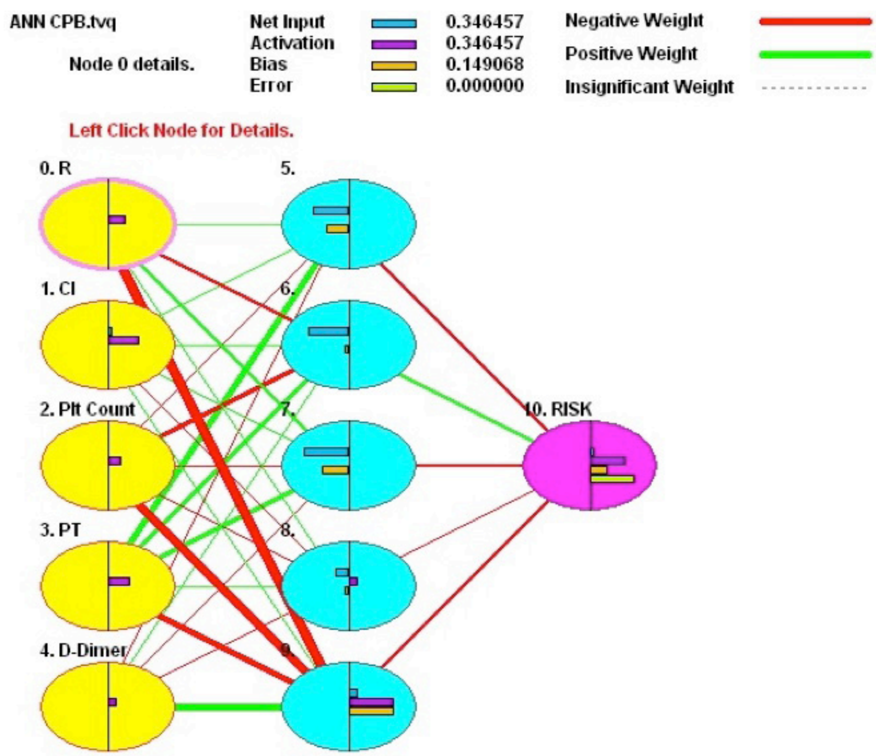


Figure 2. *Visual schematics of our neural network. First column of spheres represents the input parameters, second column of spheres represents the hidden layer, and last column of spheres represents the output parameter. Each sphere represents a node and the lines between the spheres represent the weights of the connections between the nodes.*

1. The ANN was trained with the original 15 input parameters, and the 14 parameters that have the highest relative importance were identified.
2. Subsequently, the top 14 parameters were trained and then the top 13 parameters were identified.
3. This process was repeated until the top 5 parameters were obtained.

The parameters were ranked in decreasing value of total relative importance. This is defined as the sum of relative importance's of a particular parameter in 11 iterations of subsetting as described above. Subsetting was performed to determine the relevant input parameters that would most affect the output parameter and essentially eliminates the noise in the data. It was then checked to see whether a reduction in dimensions might achieve better accuracy in prediction.

In a similar fashion as with the original 15 laboratory parameters, a new ANN with only the five most important parameters (identified by subsetting) paired with risk of bleeding was created and trained with the same 39 CPB cases. This ANN was then validated with the next 13 consecutive CPB cases and would be the one we desire to use in practice. The validation was performed by querying the ANN with the patient's input parameters in the validation set that is comprised of the next 13 consecutive cases. The patient's input parameters in the query data entry field and the ANN would give a prediction on whether the patient is expected to need a low, intermediate, or high number of blood components. The prediction is then compared to the actual usage of blood components in the case.

Results

The results of the unsupervised analysis by NeuroXL Clusterizer showed that the averages of blood usage for Clusters 1, 2, and 3 were 12.28, 14.6, and 26.8 units of blood components, respectively. The mid-points of the cluster averages were calculated and used as cut-offs for low, intermediate, and high blood usage. For example, the midpoint between clusters 1 and 2 was approximately 13.5; low blood usage was designated as <13.5 units. Using this method, the following thresholds were determined: low (≤13.5 units), intermediate (13.5-22.5 units), and high (≥22.5 units).

The subsetting results showed that the five most important parameters are Platelet, TEG-R, PT, D-Dimer, and CI, while the six least important parameters are Fibrinogen, VFN-ASA, Hgb, TEG-MA, PTT, and TT **(Table 1)**. These parameters were ranked in decreasing value of total relative importance. This is defined as the sum of relative importance's of a particular parameter in 11 iterations steps of subsetting. Of note, the validation process indicated that a five-input parameter ANN achieved a better accuracy rate of prediction compared to those with more than and less than five input parameters. Also, the top five parameters from the subsetting process were different from and yielded more accurate results when compared with the top five parameters that the ANN automatically calculates.

The validation results of the five-input parameter ANN with 13 patients in the validation set showed 9 cases with exact match, 3 cases with one-grading difference, and 1 case with two-grading difference between actual blood usage versus predicted blood usage. One-grading difference is defined as the discrepancy between low and intermediate or intermediate and high. Two-grading difference is defined as discrepancy between low and high. The results are summarized in **Table 2**.

In addition, for the ANN, convergence occurred after 2207 cycles and iterations ended with an average error rate of 0.009827. The final ANN with five input parameters was created with just 11 nodes, including five input nodes, five hidden nodes, and one output node **(Figure 2)**.

Discussion

ANN have been used in a variety of settings like pharmaceutical science, engineering, and medicine. In the medical setting, ANN have been used widely as a predictive tool in diseases such as breast cancer, gynecologic disease, lung cancer, and even sleep disorders [5,6,7,8]. In a recent study, five variables (age, surgery and radiotherapy type, tumor size, regional lymph node, distant metastasis) were used to predict the five-year survivability of breast cancer patients through the use of an ANN [5].

With good validation and convergence results, this ANN was made an accessible tool to all the pathologists and pathology residents on the hemotherapy service. The JustNN software and the data file for the ANN were downloaded onto the two Dell (Dell Inc., Round Rock, Texas) ultrabooks that are currently dedicated to the hemotherapy service. Here, the user can simply enter the five input values into the data entry fields of the query row in the ANN, and the ANN will give a prediction of whether the patient has a low, intermediate, or high risk of bleeding with associated low, intermediate, or high number of blood components used. The hemotherapy pathologist can then use this information to aid in the management of post-operative bleeding in these patients by preparing more or fewer blood products with the prediction of how much blood products the patient is likely to need.

Since the ANN became available, it has been well received by the hemotherapy physicians and residents. It is easily accessible on the Dell ultrabooks dedicated to the HT and can be downloaded from http://hemepathreview.com (section 14-Coagulation based hemotherapy) onto any Windows-based computer. The ANN has been readily used on the Hemotherapy service for the post-operative management of our patients by enabling the HT to prepare the right amount of blood products for each patient. In the future, we will be conducting a study to measure the ANN's utilization rate and how often it impacts clinical decisions.

The advantages of an ANN is that it recognizes complex nonlinear relationships, requires a smaller sample size than that used in formal statistics, has the ability to identify all possible interactions between inputs, and can be used with multiple different training algorithms [2,9]. On the other hand, the main disadvantage of ANN is that it is like a "black box". This means that once the operator creates the ANN, the software program will determine the weights of each of the input automatically [9]. Unlike in logistic regression models, the operator cannot manipulate the weight of each of the input variables and it is not possible to find causal relationship with ANN [9].

Many options were available when choosing which software to use to build this ANN. NeuroShell2 v3.0, BrainMaker v3.7, CAD/Chem v5.0, and NeuralWorks Professional II/Plus are a few examples of commercially available ANN software on the market [10]. The choice of JustNN was mainly based on the HT's previous experience with this software on another project. It is relatively easy to use, and the ANN was created using JustNN with only moderate effort.

In the end, a supervised artificial neural network (ANN) was trained to analyze intra-operative laboratory data to predict postoperative bleeding. The validation results are promising with accumulated accuracy of exact match plus one-grading difference of 92.3%. There was only a 7.7% two-grading difference. The inaccuracy in prediction in a small number of cases was most likely due to the fact that coagulation status is not the only determining factor of bleeding in patients. Laboratory tests can only measure the coagulation status of the patients. The other main factor causing post-operative bleeding in cardiac patients is surgical bleeding, referring to an anatomic defect that wasn't fully managed during surgery.

There was also good correlation between the input categories and the output categories of the final ANN, as measured by several other measures of the ANN. The learning for the ANN stops when the maximum error converges with the minimum error; for our ANN, convergence occurred after 2207 cycles. If the correlation between the input and output category had not been present in the data set, convergence would not have happened, and the training could go on indefinitely or for a very long time because the program needs to look at more patterns to find one that correlates with the data set. Hence, the longer the training goes on, the less correlation there is between the input and output parameters and it becomes a less effective prediction model. In this ANN, 2207 cycles is relatively fast, taking less than one minute to reach convergence.

Being able to estimate the amount of blood components required in the post-operative care of CPB patients based on this ANN has been of great utility. One major issue in blood banking is inventory management of blood components. For patients with a rare blood phenotype, blood components are hard to find. It takes more time to find these components. Many times they are frozen units that require thawing, and the thawed blood has a shorter shelf life. The issue here is determining how many components to prepare for the postoperative management of patients with rare blood phenotypes. If too many components are prepared, then the components may be wasted due to the shorter shelf life. On the other hand, if too few components are prepared, then there is a risk of not having the components ready when needed and patients' coagulopathy not being alleviated, with subsequent poor outcome. It allows better prediction of how many components will be needed in the post-operative care of the patient and so an optimal number of components can be prepared.

Although this is a fairly accurate prediction tool based on our validation results and our experience with this ANN, the predictions are not always on the spot. For example, there is a two-grading difference in one of the patient's actual blood usage versus the predicted blood usage in the validation set for our ANN. As discussed above, the inaccuracy is most likely due to the fact that besides coagulation status, surgical bleeding is a major factor in patients' bleeding. Another limitation of this ANN is that it only predicts total units of blood loss and does not provide a breakdown by component; however, it still does give a general sense of how many overall blood products the patient will need and is very useful in this regard.

In conclusion, to the best of our knowledge, this ANN is the first one developed for post-operative bleeding risk-stratification of CPB patients. With promising results, we have started using this ANN to risk-stratify our CPB patients. More data for the training set is being accumulated to further increase the accuracy of this ANN.

References

1. Welsh K, Nedelcu E, Bai Y, Wahed A , Klein K, Tint H, Gregoric I, Patel M, Kar B, Loyalka P, Nathan S, Loubser P, Weeks P, Radovancevic R, Nguyen A. Coagulation based hemotherapy: a novel clinical pathology consultation service for management of cardiopulmonary bypass coagulopathy. Transfusion. In press.
2. Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. J Pharm Biomed Anal. 2000;22:717-27.
3. Huang R, Nedelcu E, Bai Y, Wahed A, Klein K, Gregoric I, Patel M, Kar B, Loyalka P, Nathan S, Loubser P, Weeks P, Radovancevic R, Nguyen A. Mobile Computing Platform with Decision-support Modules for Hemotherapy. Am J Clin Pathol. 2014;141:834-40.
4. JustNN Help User Guide [internet]. Cheshire: Neural Planner Software; c2008-2014. [updated 2014 April 14; cited 2014 April 16]. Available from: http://www.justnn.com/.
5. Wang TN, Cheng CH, Chiu HW. Predicting post-treatment survivability of patients with breast cancer using Artificial Neural Network methods. Conf Proc IEEE Eng Med Biol Soc. 2013;2013:1290-1293.
6. Siristatidis CS, Chrelias C, Pouliakis A, Katsimanis E, Kassanos D. Artificial neural networks in gynaecological diseases: current and potential future applications. Med Sci Monit. 2010;16:231-6.
7. Toney LK, Vesselle HJ. Neural Networks for Nodal Staging of Non-Small Cell Lung Cancer with FDG PET and CT: Importance of Combining Uptake Values and Sizes of Nodes and Primary Tumor. 2014;270:91-8.
8. Ronzhina M, Janoušek O, Kolářová J, Nováková M, Honzík P, Provazník I. Sleep scoring using artificial neural networks. Sleep Med Rev. 2012;16:251-63.
9. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol. 1996;49:1225-31.
10. Chen Y, Jiao T, McCall TW, Baichwal AR, Meyer MC. Comparison of four artificial neural network software programs used to predict the in vitro dissolution of controlled-release tablets. Pharm Dev Technol. 2002;7:373-9.