Received: 2 August 2023 | Revised: 25 September 2023 | Accepted: 27 September 2023

DOI: 10.1111/trf.17582

REVIEW ARTICLE

TRANSFUSION

Machine learning in transfusion medicine: A scoping review

Suzanne Maynard[1,2,3] | **Joseph Farrington**[4] | **Samah Alimam**[5] | **Hayley Evans**[2] | **Kezhi Li**[4] | **Wai Keong Wong**[6] | **Simon J. Stanworth**[1,2,3]

[1]Medical Sciences Division, Radcliffe Department of Medicine, University of Oxford, Oxford, UK

[2]NIHR Blood and Transplant Research Unit in Data Driven Transfusion Practice, Nuffield Division of Clinical Laboratory Sciences, Radcliffe Department of Medicine, University of Oxford, Oxford, UK

[3]NHSBT and Oxford University Hospitals NHS Foundation Trust, Oxford, UK

[4]Institute of Health Informatics, University College London, London, UK

[5]Haematology Department, University College London Hospitals NHS Foundation Trust, London, UK

[6]Director of Digital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Correspondence
Suzanne Maynard, NHS Blood and Transplant, John Radcliffe Hospital, Headley Way, Headington, Oxford OX3 9BQ, UK.
Email: suzanne.maynard@nhs.net

Funding information
NIHR Blood and Transplant Research Unit in Data Driven Transfusion Practice, Grant/Award Number: NIHR203334; UK Research and Innovation Centre for Doctoral Training (CDT) in AI-enabled Healthcare Systems, Grant/Award Number: EP/S021612/1; NIHR University College London Hospitals Biomedical Research Centre

KEYWORDS: artificial intelligence, machine learning, patient blood management, transfusion

1 | INTRODUCTION

Blood transfusion is a routine medical procedure in hospitals with over 2 million blood products transfused in the UK every year at a cost of over £300 million and a median national rate of 34 packed red cells per 1000 population in Europe.[1,2] A blood transfusion can be life-saving but can also cause harm.[3] Repeated studies have demonstrated a gap between recommended blood use and clinical practice.[4,5] National challenges with blood stock shortages highlight the need to optimize our current approach to identify who requires and benefits from blood components.[6]

Recent advances in digital technology offer a wealth of new tools, which can help improve clinical practice as well as improving both the equality and equity of healthcare. Patient Public Involvement groups consistently support better use of data and better understanding of how it might improve efficiencies, prioritizing the need for healthcare professionals to engage with research optimizing use of data. Machine learning (ML) is a subfield of artificial intelligence (AI), which offers the ability to integrate complex and varied data types and could support clinician decision-making, aid personalized care, and, with additional work, improve patient outcomes.[7,8] This field is a rapidly advancing one, which has the potential to revolutionize patient blood management (PBM).

Abbreviations: AI, artificial intelligence; ATR, acute transfusion reactions; AUROC, area under the receiver operating characteristic curve; EHRs, electronic health records; ICU, intensive care unit; LR, logistic regression; ML, machine learning; MTP, massive transfusion protocol; NLP, natural language processing; NSQIP, National Surgical Quality Improvement Program; PBM, patient blood management; PRBCs, packed red blood cells; SHAP, SHapley Additive exPlanation; TRALI, transfusion-associated lung injury.

Suzanne Maynard and Joseph Farrington contributed equally to this study.

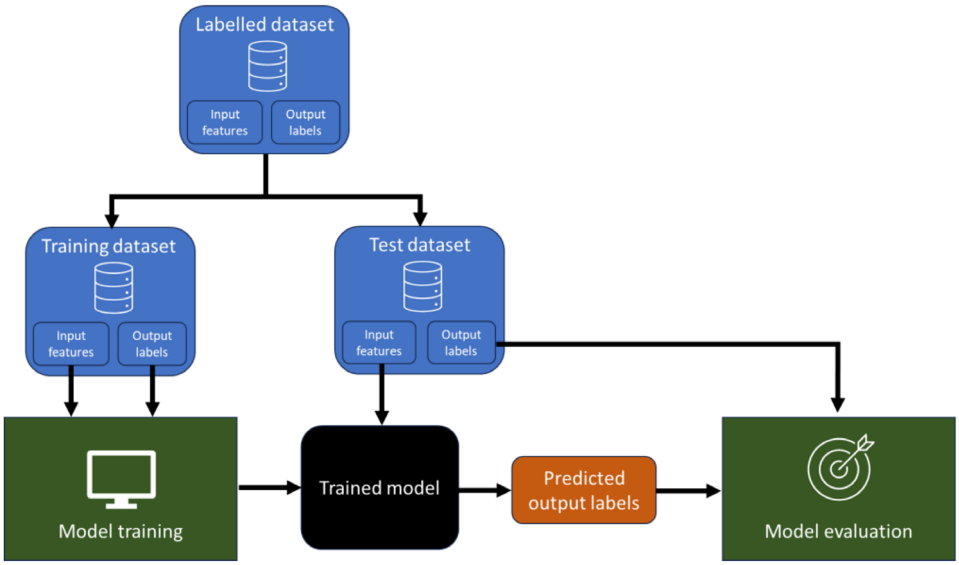

FIGURE 1 A typical supervised machine learning workflow. Supervised learning requires a labeled dataset in which we already know the outcome we are trying to predict, the output label, for each example. To estimate how well the model will perform on future, unseen, examples, we hold back a portion of the data (the test set) from the training process and use this dataset to evaluate the model's performance. If the test dataset comes from the same site(s) as the training dataset, then it may be described as an internal validation set. If the test dataset comes from one or more different sites or settings, then it may be described as an external validation set. [Color figure can be viewed at wileyonlinelibrary.com]

Successful implementation of ML to support clinical workflows requires collaboration between computer scientists and clinicians. Key features of ML have been described elsewhere, which support informed interpretation of the literature.[9,10] The majority of work applying ML to healthcare uses supervised learning whereby the model is trained on input features and labeled output features to enable predictions on unseen examples (Figure 1).[11] Model performance evaluation uses metrics, which summarize prediction quality, for example, the area under the receiver operating characteristic curve (AUROC) for classification models. The two other main categories of ML approaches are unsupervised learning, which identifies patterns in unlabeled data (e.g., finding clusters of similar patients), and reinforcement learning, an approach to learning how to act through trial and error. To be useful in practice, models need to be validated and integrated into a clinical workflow, where capacity constraints and users ignoring alerts may limit the impact of even a perfectly performing model.[12]

The purpose of this review is to collate the breadth of literature of ML in transfusion medicine, describing current trends and capturing key methodological approaches, adding to the recognized need for up-to-date discussion of the challenges and potential solutions to the prospective implementation of ML in transfusion medicine.[13]

2 | METHODS

2.1 | Scope of the review

The review aimed to report on original research articles, using ML approaches with a focus on transfusion medicine. We followed the approaches of a scoping review used by Cochrane from the Canadian Institutes of Health Research, defined as "exploratory projects that systematically map the literature available on a topic, identifying key concepts, theories, sources of evidence, and gaps in the research."[14]

2.2 | Eligibility criteria

Eligibility for studies was defined by blood transfusion in humans (or the support of transfusion) as the main outcome. There were no restrictions on year of publication, publication status, or language. We excluded studies using linear or logistic regression (LR) primarily for statistical inference and/or to construct a predictive risk score. This exclusion is consistent with a recent systematic review on the impact of ML on patient care.[15] When considering inventory management in a hospital blood bank, we focused on recent work using patient data from electronic health records and, therefore, excluded

research that predicted future demand based solely on historic demand. As it is common in ML and computer science to submit full-length works to top tier conferences instead of journals, where reports met all other criteria, conference articles were included as full text ($n = 3$).

2.3 | Search strategy

We searched the Clarivate Web of Science database on January 4, 2023 with the following search terms: [TS = (machine learning OR artificial intelligence OR forecast* OR algorithm OR prediction model OR predictive model OR neural network)] AND TS = {transfus* OR blood product OR blood bank OR [reaction NEAR (blood OR transfus*)]}. We reviewed the lists of publications in the literature and consulted with all authors. Additional citation search was performed and relevant reports were added, which were not captured in the initial Web of Science search.

2.4 | Data extraction

The title and abstract screen were conducted in duplicate (SM and JF). Differences were resolved by consensus or with a third reviewer, to arrive at the final set for full-text review. Where the same work was published in more than one journal (i.e., not a true duplicate, but papers aimed at different audiences), we selected the journal with a medical focus where possible for inclusion in the summary tables and figures. Data were extracted in duplicate (SM and JF), with discrepancies resolved by a third reviewer.

2.5 | Data analysis

Results were presented descriptively. Initial clinical categories were defined and agreed based on understanding of the literature and were further refined following title and abstract screening. We extracted information on clinical applications, data sources, and ML methods. The research team also predefined a range of factors identified as important when considering the methodology and exploring the opportunities and limitations of translating ML models to a healthcare setting.[16–18] Meta-analysis of the results was not undertaken due to the wide range of different tasks, variability in definitions for similar tasks, and reporting heterogeneity.

3 | RESULTS

3.1 | Search selection

A total of 4504 publications were retrieved using the described search strategy performed on January 4, 2023 (Figure 2). Initial screening returned 107 citations, and 93 articles were selected for inclusion in the study following full-text review, including the addition of two articles identified through citation searching. Overall, 16 studies eligible for full-text review were excluded: Three were duplicates captured through alternative journal publications, three did not meet ML criteria, transfusion was not the main outcome for seven, and the full-text article was not available for two studies. One article was removed due to subsequent publication retraction.

3.2 | Temporal distribution and study categories

The temporal distribution of 93 included publications is shown in Figure 3. There is a clear trend toward increasing frequency of publications over time with 56% (52/93) of the articles published in the last 3 years.

3.3 | Clinical setting

The majority of studies were focused on prediction of transfusion (58%) with other key areas of ML application identified within transfusion safety (22%), hospital blood bank (10%), and supporting transfusion decisions (10%) (Figure 4A). Within prediction of transfusion (Figure 4B), a significant majority of studies were in the setting of surgery (61% 33/54), followed by trauma (24% 13/54). In the remaining eight studies, ML was deployed in the setting of obstetrics, gastrointestinal bleeding, and hemato/oncology, and in three studies applied more broadly to all inpatients and intensive care, captured as "other hospital settings."

3.4 | Study objectives and findings

The objectives, sample size, and key findings of all studies within these broad categories of clinical settings are provided in Table 1 and more detailed methodological considerations in Table 2. Overall clinical applications, trends, and a summary of main findings are discussed in more detail under the relevant subheading below.

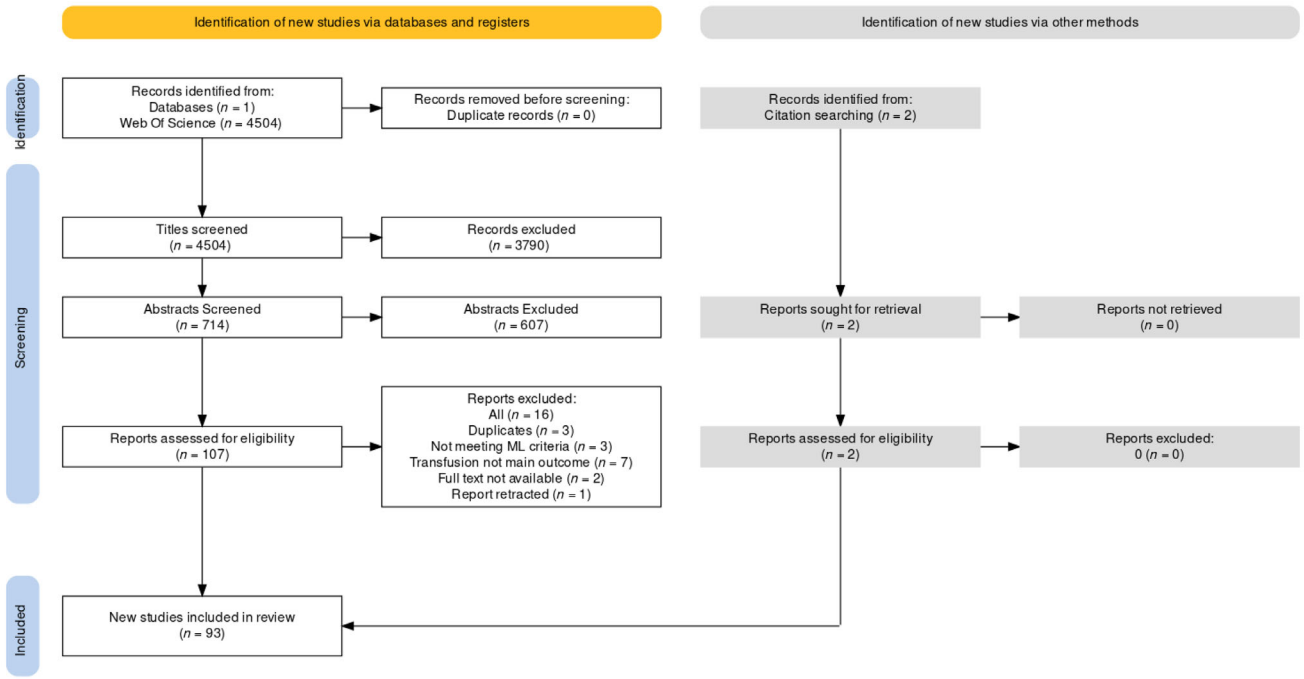

FIGURE 2 Modified Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart of the search results.[112] [Color figure can be viewed at wileyonlinelibrary.com]

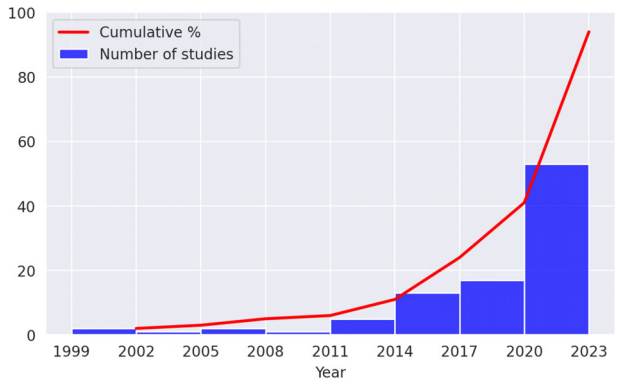

FIGURE 3 Histogram with cumulative frequency line of papers by publication year. [Color figure can be viewed at wileyonlinelibrary.com]

3.5 | Baseline features

Overall, the most common countries for identified studies are the United States (44), followed by China (16), Europe (12), and Canada (6). The range of sample sizes reported in the studies varied from 41 to more than 4 million (Table 1). Packed red blood cells (PRBCs) were the focus of approximately half of the studies (46/93), with most of the remainder considering either multiple blood products (22%) or not specifying the blood products (22%). Five studies (5%) considered only platelets, and two studies (2%) considered only plasma.

3.6 | Predicting transfusion

3.6.1 | Surgery

The majority of identified studies employed ML to predict transfusion related to a specific specialty or procedure, notably within orthopedics,[19–25] cardiac surgery,[26–30] spinal surgery,[31–34] and liver transplant,[35–37] focusing on a specific procedure or a variety within that specialty (Table 2). A small number of studies consider procedures from multiple specialties[38–44] with Walczak and Velanovich[43] including 56 different surgeries from the publicly available United States National Surgical Quality Improvement Program (NSQIP). Their use of single models to predict transfusions for a wide variety of surgical procedures could provide a much simpler approach rather than individual models for each surgical procedure.

Some researchers interrogated models to identify features to help predict PRBC transfusion[22,32,34,36] or the decision to transfuse[44] as examples of hypothesis generation from ML. Five studies developed online risk calculators and web apps based on their models.[24,30,32,45,46] Gurm et al.[30] highlighted that previous simplified non-computerized tools need no longer be the limit to what can be utilized in clinical medicine; however, a recent systematic review concluded that the resultant clinical prediction models for blood transfusion in elective surgery are of a high risk of bias and often fail to adhere to

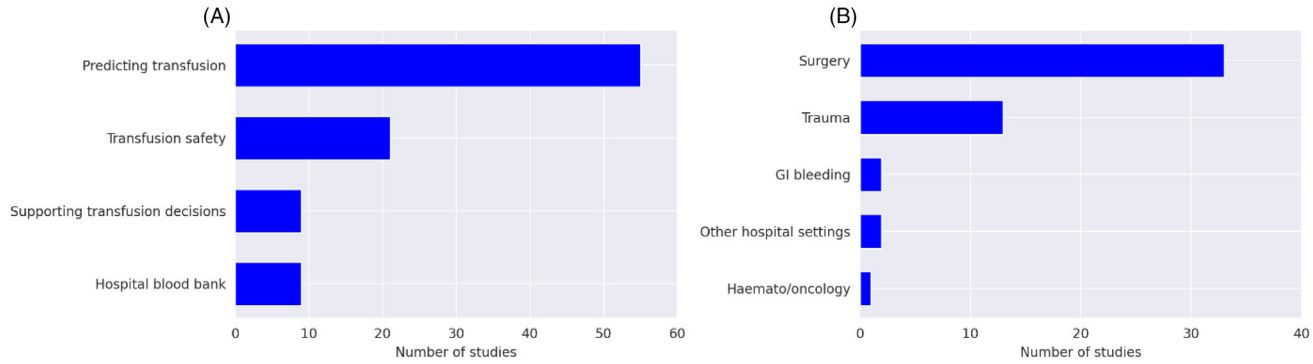

FIGURE 4 Bar charts of papers included after full-text review: (A) split by group and (B) split by subgroup within predicting transfusion. [Color figure can be viewed at wileyonlinelibrary.com]

reporting standards, emphasizing caution before application to clinical practice.[47]

3.6.2 | Trauma

There is an extensive body of literature developing risk scores for transfusion in trauma patients, and multiple reviews suggest further model development and/or validation is required.[48–51] Two key challenges with trauma are the potentially large requirements of blood for a small proportion of patients[52] and the importance of a fast response.[53] The activation of the massive transfusion protocol (MTP) is resource intensive and may result in product wastage in cases of false positive activation.[54] The ability to predict future transfusion requirements prior to hospital arrival can support triage decisions and help to ensure that blood products are available when required on arrival.[55–57] When making predictions using data collected at the hospital, research has focused on four related prediction tasks: predicting transfusion,[58,59] the number of units transfused,[60] activation of the MTP,[54] and/or massive transfusion.[58,61–63]

The model developed by Mina et al.[54] to predict MTP activation was integrated into a smartphone application, externally validated and an implementation and prospective validation study was conducted at the initial site. Clinicians informed of the model's prediction made better decisions in the prospective validation study.[54,64,65] This is a key demonstration of how we expect such models will eventually be used in practice: supplementing rather than replacing clinical judgment.

3.6.3 | Obstetrics, gastrointestinal bleeding, and hemato/oncology

Demand for blood components and associated morbidity and mortality are significant in obstetrics, gastrointestinal bleeding, and hemato/oncology[1,5,66]; however, ML for prediction of transfusion in these settings is underrepresented comprising a total of 5 of 54 studies, none of which have undergone prospective validation or implementation at the time of writing. The studies exploring gastrointestinal bleeding demonstrate benefits of using large, publicly available data sets, able to externally validate models.[67,68] Given the availability of the data, these tasks could be developed into benchmarks, enabling different research teams to compare the performance of new approaches. Interestingly, Levi et al.[67] apply their model to support triage: predicting which patients do not require transfusion (suggesting no ongoing bleeding) and, therefore, may avoid admission to intensive care unit (ICU). Shung et al.[68] highlight the potential impact of alert fatigue in the context of repeated predictions on a problem with relatively low frequency.

3.6.4 | Other hospital settings

Lee et al.[69] and Ghassemi et al.[70] predict blood transfusion within the ICU, respectively, demonstrating the inadequacy of hemoglobin measurement alone as a determinate of transfusion and that general patient state representations could be used to better predict platelet and plasma transfusions.

3.6.5 | Findings across all studies predicting transfusion

Review of all studies suggested that task-specific performance of ML for predicting transfusion need is frequently reported with AUROC >0.8 (Table 1). In 13 studies that reported a direct task-matched comparison of ML to LR models, LR matched or outperformed ML in 54% (Table 1). However, in additional seven studies, ML was reported to demonstrate measurable clinical

TABLE 1 Summary of objectives and key findings of machine learning (ML) in transfusion.

Reference	Objectives as stated	Sample size *n*	Key findings as reported
Predicting transfusion			
GI bleeding			
Levi et al. (2021)[67]	Predict transfusions in patients admitted to ICU with GI bleed as surrogate for rebleeding	14,853	Ensemble method achieved AUROC greater than 0.8 and achieved similar performance on both ICU databases
Shung et al. (2021)[68]	Predict need for PRBC transfusion within 24 h from ICU admission for GI bleeding	4,050	LSTM performed best, AUROC 0.81 on internal validation and 0.65 on external validation set
Hemato/oncology			
Ho and Chang (2011)[113]	Predict platelet requirements in acute myeloid leukemia	744	FCNN trained with genetic algorithm outperforms FCNN trained using conventional backpropagation (accuracy 78.1%–88.6% vs. 53.3%–68.2%, respectively)
Obstetrics			
Aldaghi et al. (2020)[114]	Data mining to develop a new system for inpatient preoperative ordering of PRBCs in obstetrics and gynecology	1,097	DT achieved accuracy 96.1%, cross-matched units reduced by 71.5% compared to historical practice
Pressly et al. (2021)[110]	Predict transfusion risk anti-peri- and postpartum in obstetrics	63,973	ML did not improve on logistic regression (LR). Need for improved data collection and curation in obstetrics.
Other hospital settings			
Ghassemi et al. (2017)[70]	Use ICU patient state to predict the onset of interventions including PRBC, plt, and plasma transfusion	36,050	Features from unsupervised ML improved classification performance for predicting all interventions except PRBC transfusion
Lee et al. (2022)[69]	Predict PRBC transfusion for ICU patients within 24 h of admission	16,222	A LightGBM model achieved the best AUROC 0.91
Mitterecker et al. (2020)[115]	Predict transfusion requirements for all hospital admissions	206,271	LR and ML models performed similarly, AUROC 0.97 for predicting PRBC transfusion
Surgery (Orthopedic)			
Chang et al. (2018)[19]	Predict transfusion requirements within 48 h of orthopedic surgery	1,396	LR outperformed ML models (AUROC 0.79)
Cohen-Levy et al. (2022)[20]	Predict transfusion rates following primary total hip arthroplasty	7,265	FCNN achieved the best predictive performance (AUROC 0.82)
Gowd et al. (2019)[21]	Demonstrate that supervised ML models can better predict postoperative complications after total shoulder arthroplasty (TSA) than comorbidity indices	17,119	GBM and LR both achieved AUROC 0.77, compared to AUROC 0.64 for baseline model using ASA classification
Huang et al. (2018)[22]	Identify the predictors of need for blood transfusion in primary lower limb total joint arthroplasty	15,187	Risk factors identified and RF outperformed LR (AUROC 0.84 vs. 0.77)

TABLE 1 (Continued)

Reference	Objectives as stated	Sample size *n*	Key findings as reported
Huang et al. (2021)[23]	Identify predictors of postoperative transfusion in elective hip and knee total joint arthroplasty and compare performance of ML predicting transfusion	12,642	RF achieved the highest AUROC (0.86) compared to 0.75 for LR
Huang et al. (2021)[102]	Perioperative PRBC transfusion prediction model in pelvic surgery	510	XGBoost achieved the best predictive performance. The model outperformed blood preparation based on surgeons' experience. Prospective accuracy 81.8%.
Jo et al. (2020)[24]	Identify preoperative variables to predict transfusion after total knee arthroplasty	2,096	GBM achieved an AUROC of 0.842 on the internal validation and 0.880 on external validation
Mohammed et al. (2022)[25]	Prediction of blood transfusion after TKA	636,062	All ML models performed better than LR. GBM performed best, with AUROC 0.797.
Surgery (cardiac)			
Cevenini et al. (2013)[26]	Develop a locally customized model for planning transfusion requirements in cardiac surgery	3,182	Binary classification using naïve Bayes achieved 71.2% sensitivity and 78.4% specificity but multiclass performance was poor
Gurm et al. (2014)[30]	Predicting blood transfusion in patients undergoing contemporary PCI	103,294	Random forest model achieved AUROC 0.89 with all features and 0.88 with fewer variables for ease of use
Liu et al. (2021)[27]	Predict PRBC transfusion during mitral valve surgery	698	CatBoost achieved the highest AUROC, at 0.888
Tschoellitsch et al. (2022)[28]	Predict massive perioperative allogenic blood transfusion in cardiac surgery	3,782	RF achieved AUROC 0.81 and outperformed predictions of expert clinicians
Wang et al. (2022)[29]	Use preoperative variables to predict intraoperative blood use in cardiothoracic surgery	2,847	Final model combines Gaussian process classification for >3 PRBCs (AUROC 0.826) and regression 0–3 PRBCs (RMSE 0.117 for 0 units, 1.705 for 1–3 units)
Surgery (spinal)			
Dong et al. (2021)[32]	Develop a prediction model to evaluate blood transfusion risk after spinal fusion for spinal tuberculosis	152	Nomogram based on LR achieved best performance (AUROC 0.75)
Durand et al. (2018)[33]	Predict blood transfusion during, or in the first 72 h after, ASD surgery	1,029	CART model selected due to ease of implementation despite and AUROC not significantly lower than with RF (0.79 vs. 0.85)
Raman et al. (2020)[34]	Predict variables, which best predict perioperative blood transfusion requirements in ASD surgery	909	ML identified variables associated with perioperative PRBC transfusion
Ramos et al. (2022)[31]	Predict the need to PRBC transfusion after ASD surgery	1,173	The best FCNN model achieved an AUROC of 0.84

(Continues)

TABLE 1 (Continued)

Reference	Objectives as stated	Sample size *n*	Key findings as reported
Surgery (liver transplant)			
Chen et al. (2022)[35]	Predict intraoperative massive blood transfusion in liver transplantation	1,239	CatBoost performed best on internal and prospective validation data (AUROC 0.81 and 0.75, respectively)
Cywinski et al. (2014)[36]	Determine patient variables to predict recipient risk for large blood transfusion requirements during OLT	804	Models were unable to reliably predict which patients might require the largest amount of blood during OLT
Liu et al. (2021)[37]	Identify preoperative variables and predict PRBC transfusion during or after liver transplant	1,193	XGBoost achieved best performance (AUROC 0.813) and performed well on prospective data (accuracy 76.9%)
Surgery (multiple specialties)			
Etchells et al. (2006)[44]	Determine the rules that staff of a department of anesthesia use when deciding to transfuse a patient perioperatively	209	Rule extraction summarized decision-making strategies of clinicians reproduced transfusion decisions with specificity 0.96 and sensitivity 0.93
Feng et al. (2021)[38]	Predict perioperative PRBC transfusion requirements from information available preoperatively	130,996	LightGBM outperformed LR (AUROC 0.908 vs. 0.868) and clinicians in predicting if a transfusion would be required
Hayn et al. (2017)[40]	Predict transfusion in elective surgery for data-driven benchmarking of transfusion patterns	6,530	RF predicts PRBC transfusion volume with correlation coefficient 0.61. Significant differences in feature importance between hospitals.
Hayn et al. (2016)[39]	Predict PRBC transfusion requirements during surgery based on features available at different time points	6,530	Use of features available before surgery versus historic data improves performance for blood transfusion ($r = 0.45$ vs. 0.20)
Lou et al. (2022)[41]	Predict likelihood of PRBC transfusion using surgery-specific and patient-specific variables	4,142,111	GBM performed best achieving PPV of 0.06 and 0.21 at 96% sensitivity on internal and external validation respectively
Ngufor et al. (2015)[42]	Predict intraoperative PRBC transfusion, and other surgical outcomes	1,234	Multitask outperformed single-task (AUROC 0.86 vs. 0.82) for intraoperative blood transfusion
Walczak and Velanovich (2020)[43]	Predict PRBC and whole blood transfusions for all inpatient operations	1,337,777	FCNNs can predict >75% of patients who will require transfusion and 70% of those who will not (AUROC 0.85–0.86)
Surgery (other)			
Jalali et al. (2021)[45]	Predict blood transfusion requirements for pediatric craniofacial surgery	2,143	GBM performed best at both predicting transfusion (AUROC 0.87) and number of units ordered preoperatively (R^2 0.73)
Tunthanathip et al. (2022)[46]	Identify cost differences between strategies for preoperative blood product preparation for patients with brain tumors	1,681	ML-based strategy predicted cost savings of USD 93,000 (67.88%) compared to routine hospital preparation

TABLE 1 (Continued)

Reference	Objectives as stated	Sample size *n*	Key findings as reported
Walczak et al. (2000)[116]	Predict blood transfusion requirements for patients undergoing abdominal surgery	220 (AAA) c.175 (non-AAA)	Lower CTR was achieved in three of four procedures using FCNN compared to MSBOS; however, FCNN also under-predicted in individual patients
Walczak et al. (2000)[117]	Predict the number of PRBCs required for transfusion during AAA operations	109	FCNN saved 16 units below MSBOS and C/T ratio of 2.71—potential annual savings of over USD 1000 for AAA operations
Walczak et al. (2021)[118]	Predict transfusion requirements (and estimated blood loss) during or in the first 24 h after myomectomy	96	FCNN predicts transfusion with sensitivity 71.4% and specificity 85.4%
Zhang et al. (2022)[119]	ML modified nomogram to predict perioperative transfusion in total gastrectomy	513	LR outperformed ML with AUROC 0.85 used to build nomogram (C index 0.83)
Trauma			
Avital et al. (2021)[55]	Identify factors associated with emergency department transfusion and use to predict need	2,885	LR achieved AUROC of 0.82, similar to or better than ML approaches
Christie et al. (2019)[58]	Predict need for transfusion after severe injury	1,494	Ensemble method achieved overall AUROC of 0.87–0.90 across multiple post-injury timepoints
Dente et al. (2021)[65]	Implementation of model to predict activation of massive transfusion protocol (MTP)	321	Surgeons' decision to activate MTP improved with information from model prediction, from AUROC of 0.86–0.93
Feng et al. (2021)[59]	Assist doctors in quickly making decisions about need for blood transfusion after trauma	1,371	XGBoost performs best with all features (AUROC 0.94), but LR best with readily available feature subset (AUROC 0.72)
Hodgman et al. (2018)[64]	External validation of model to predict activation of the MTP	1,245	AUROC 0.694 for predicting MTP activation and AUROC 0.695–0.711 for predicting MT administration compared to 0.96 in original study
Lammers et al. (2022)[120]	Predict patients at the highest risk for transfusion on the battlefield	22,158	RF achieved the best predictive performance, AUROC of 0.98
Feng et al. (2021)[59]	Predict high MT risk to improve clinical decision-making in multiple trauma	478	Decision trees help identify risk factors and achieved AUROC of 0.85
Liu et al. (2015)[56]	Prehospital vital signs to identify patients with hemorrhagic injury who receive PRBC transfusion in first 24 h	855	Sensitivity 76% for PRBC transfusion within 24 h of admission, with specificity 87% for no transfusion. Prospective performance similar to retrospective
Mina et al. (2013)[54]	Predict MT of critically injured patients	10,900	Mean AUROC of 0.96 and embedded into smartphone app
Nederpelt et al. (2021)[57]	Predict the need for early MT in truncal gunshot wounds to support in-field triage	29,816	AUROC 0.86 for prediction of early massive transfusion, and correct predictions assigned high confidence by separate neural network

(Continues)

TABLE 1 (Continued)

Reference	Objectives as stated	Sample size *n*	Key findings as reported
Seheult et al. (2019)[61]	Predict MT within 24 h of hospital admission due to trauma	825	Decision tree performed better compared to previously established scoring systems, ABC score and TASH
Shahi et al. (2021)[62]	Predict MT in pediatric blunt solid organ injury to help decision-making	477	Accuracy of prediction of massive transfusion 90.5% and outperformed existing approaches for identifying emergent intervention such as ABCD
Walczak et al. (2005)[60]	Predict transfusion for trauma patients using features available on entry to ED	1,016	Performance of FCNN to predict transfusion meets or exceeds LR
Supporting transfusion decisions			
Adverse outcome			
Bruun-Rasmussen et al. (2022)[75]	Investigate the effect of PRBC donor sex on recipient mortality	90,917	ML to emulate a randomized control trial identifies sex-matched transfusion policy may benefit patients
Ngufor et al. (2015)[73]	Study the effect of preoperative plasma transfusion on perioperative bleeding	14,743	Perioperative plasma transfusion increases risk of perioperative bleeding and intraoperative PRBC transfusion
Ngufor et al. (2018)[76]	Identify subgroups with differential effects of plasma transfusion using unsupervised learning	3,135	Unsupervised RF identified clinically meaningful subgroups regarding effect of plasma transfusion on bleeding and mortality
Nguyen et al. (2020)[74]	Augment analysis of PROPPR trial to determine impact of transfusion ratios on patient outcomes in trauma	680	Transfusion ratios did not significantly affect mortality but did affect hemostasis
Appropriateness			
Barbosa et al. (2011)[79]	Identify variables in trauma indicating futility of MT (measured by mortality at 30 days)	704	No variable combinations identified by ML to predict futility except patients over 65 with severe head injuries
McGlothlin et al. (2017)[78]	Retrospectively identify clinically inappropriate transfusions	Not stated	Implementation of blood utilization dashboard reduced inappropriate red cell transfusion by 10.2%
Yao et al. (2019)[77]	Efficiently check the appropriateness of blood transfusion on a large volume of cases using ML	4,946	ML achieved a 96.8% overall match rate with human experts
Dose			
Epah et al. (2022)[71]	Predict hemoglobin/iron content of any given PRBC unit from routinely collected data	8,695	Linear regression preferred. $R^2 > 0.9$ for predicting both Hb and iron.
Zhang et al. (2019)[72]	Rapid and nondestructive quantitative analysis of bagged liquid products	59	Improved the analysis accuracy of Hb in blood bags, correlation coefficient of Hb 0.9915

TABLE 1 (Continued)

Reference	Objectives as stated	Sample size *n*	Key findings as reported
Transfusion safety			
Hemovigilance			
Clifford et al. (2013)[82]	EHR-based screening algorithms for improved detection of TRALI/transfused acute lung injury (tALI) and TACO	223	ML may improve case identification because only 14% of TRALI/tALI and 11% of TACO true positives identified by the algorithm were reported
Kim et al. (2015)[121]	External validation of a previously developed CART model (Clifford et al, 2013[82]) to identify TACO and TRALI	500	Limited diagnostic value due to the low specificity for detecting TRALI (57%) and TACO (56%)
Murphree et al. (2015)[122]	Predict the risk of TACO and TRALI for noncardiac surgery patients prior to transfusion	3,255	AUROC for final models ranges between 0.72 and 0.84, with a LR model achieving sensitivity of up to 93%
Murphree et al. (2015)[123]	Deploy previously developed models (Murphree 2015) for predicting TACO and TRALI into a perioperative health information system	N/A—not developing models	Prototype model producing live alerts for review by development team
Murphree et al. (2015)[124]	Ensemble learning to predict TACO and TRALI pre-transfusion for noncardiac surgery patients	3,398	Ensemble models did not outperform the base models, and the mean AUROC of the top five models was 0.84
Nguyen et al. (2014)[84]	Predict risk of acute transfusion reaction (ATRs) following platelet transfusion modeled on biologic response modifiers	124	Decision tree models can be used to understand the relationship between biologic response modifiers and ATRs
Ramoa et al. (2021)[125]	Develop a computer support system to automatically classify and validate ATRs	3,829	Ensemble classifier correctly predicted 823 of 959 reaction types. Error analysis highlighted deficiencies in current reporting system.
Roubinian et al. (2020)[83]	Investigate whether NT-proBNP has utility in the identification and classification of pulmonary transfusion reactions	495	CART models using NT-proBNP (AUROC 0.83) achieved similar performance to model using echocardiogram
Torres et al. (2014)[85]	Identify determining factors associated with TRALI using an evolutionary algorithm	174	Previous risk factors including age and the ratio PaO_2/FiO_2 confirmed, and identified a novel link to smoking that surprised physicians
Tsatsoulis et al. (2003)[80]	Identify clusters of similar adverse transfusion events in large databases to help determine patterns, trends, and best practices	c. 600 reports	Combining case-based reasoning and information retrieval produces clusters that better match expert judgment compared to either approach alone
Whitaker et al. (2022)[81]	Explore whether machine learning methods, such as NLP, can identify and report transfusion allergic reactions from EHR	751 events/443 patients	Model including NLP features achieved an AUROC of 0.92, versus 0.89 for a model without NLP features

(Continues)

TABLE 1 (Continued)

Reference	Objectives as stated	Sample size *n*	Key findings as reported
Yamada et al. (2022)[86]	Identify factors associated with transfusion-associated hyperkalemia (TAH) in children	95	The biggest factor in TAH occurrence was the total transfused volume within 12 h, followed by the age of PRBC units
Zhu et al. (2022)[87]	Understand risk factors for adverse events during exchange transfusion in severe neonatal hyperbilirubinemia	188	XGBoost model achieved a mean AUROC of 0.71 and revealed nonlinear relationships potentially relevant for practice
Laboratory			
Borgmann et al. (2016)[126]	Identify red cell RhD phenotypes, including low expressed variants, to avoid alloimmunization	51 blood samples	Ensemble model achieves classification accuracy of 96%, minimum recall 92% for low expressed DEL phenotype
Darlis et al. (2018)[127]	Hardware implementation of a neural networks model to identify human blood type from grayscale images	Unclear	The models achieved classification accuracy of up to 97.5%
Doan et al. (2020)[91]	Nonsubjective assessment of blood storage lesions to predict product quality	38 bags/∼67,500 cells	Model trained with storage duration instead of expert assessment learns features that predict blood quality better than morphological assessment
Ferraz et al. (2017)[128]	Determine blood type compatibilities of an individual in an emergency	41 patients	Both AdaBoost and SVM models are able to achieve an F1-score of 1
Kim et al. (2022)[92]	Novel approach for fast phenotypic assessment of PRBC storage lesions	3 donors; 219 images	Deep learning achieved classification accuracy of 95% with a high throughput (152 cells/s)
Larpant et al. (2022)[90]	Phenotyping red cell antigens with deep learning-assisted decision-making	4,692	Deep learning-assisted human decision-making resulted in classification of ambiguous Rh antigens with 100% accuracy
Rosales et al. (2022)[129]	Automatically identify individual blood type using image processing and machine learning algorithms	Unclear	Coarse tree DT outperformed LR and other ML (performance accuracy 97.77%)
Wu et al. (2022)[89]	Deep learning for the automated classification of incomplete antibody reaction intensity (IARI) to support hemolytic disease screening	1,628	Ensemble model achieved accuracy of 99.8%, immunologist accuracy improved with model use (average +6.1%)
Hospital blood bank inventory management			
Ordering			
Fischer et al. (2010)[99]	Predict blood requirements to minimize hospital storage costs and low stocks	60,784 and 59,968	Better performance when predicting blood requirements per medical activity than when predicting aggregate demand
Guan et al. (2017)[94]	Forecast future platelet demand to guide collection from donors and reduce wastage	30,000	ML model reduced wastage from 10.5% to 3.2% in simulation
Li et al. (2022)[96]	Develop a data-driven demand forecasting and inventory	227,944	The STL-XGBoost model reduced costs by 43% without stock shortages in simulation

TABLE 1 (Continued)

Reference	Objectives as stated	Sample size *n*	Key findings as reported
	management strategy for red blood cells		
Mirjalili et al. (2022)[95]	Predict platelet demand for the following day with simulator to evaluate potential to reduce wastage and shortages	6,220	ML model reduced wastage from 4.01% to 2.54% and shortage from 0.44% to 0.05% in simulation
Quinn et al. (2019)[93]	Recommend PRBC order quantity based on admitted patient characteristics and historic demand	26,326	Significant reduction in PRBC outdate rates (1.72%–0.72%) and daily inventory levels in 12 months after implementation compared to preceding 12 months
Schilling et al. (2022)[97]	Forecast platelet demand and calculate possible waste reduction and platelet shortage	45,900	ML models reduced both shortages (6.5%–2%) and expiries (10%–5%) in simulation
Sun et al. (2021)[98]	Predict red blood cell demand on the following day for each ABO blood group	1,243 days (transfusions unclear)	The XGBoost model achieved the best performance with a mean absolute error between 5.91 and 11.19
Waste			
Rad et al. (2020)[100]	Predict blood product discard to reduce wastage and optimize inventory management	17,108	All-K-order-Markov model achieved 79% accuracy in predicting discards. Visualizations can help laboratory managers investigate the causes of discards.
Xiang et al. (2021)[101]	Identify patterns associated with blood product wastage	879,532	Wastage factors included specific wards, smaller blood banks, and evening work shifts

Abbreviations: AAA, abdominal aortic aneurysm; ASA, American Society of Anesthesiologists; ASD, adult spinal deformity; ATR, acute transfusion reaction; AUROC, area under the receiver operating curve; CART, classification and regression tree; CTR, cross-match to transfusion ratio; DT, decision tree; ED, emergency department; FCNN, fully-connected neural network; FiO_2, fraction of inspired oxygen; GBM, gradient boosting machine; GI, gastrointestinal; ICU, intensive care unit; LR, logistic regression; ML, machine learning; MSBOS, maximal surgical blood ordering schedule; MT, massive transfusion; MTP, massive transfusion protocol; NLP, natural language processing; NT-proBNP, N-terminal pro-brain natriuretic peptide; OLT, orthotopic liver transplant; PaO_2, partial pressure of oxygen; PBM, patient blood management; PCI, percutaneous coronary intervention; plt, platelets; PPV, positive predictive value; PRBC, packed red bloods cells; PROPPR, pragmatic, randomized optimal platelet, and plasma ratios; RF, random forest; RMSE, root mean squared error; STL, seasonal and trend decomposition using loess; SVM, support vector machine; TACO, transfusion-associated circulatory overload; TAH, transfusion-associated hyperkalemia; TASH, trauma-associated severe hemorrhage; TKA, total knee replacement; TRALI, transfusion-associated lung injury; USD, United States dollar; XGBoost, eXtreme gradient boosting.

improvements such as cost savings or performance over current scoring systems (Table 1).

3.7 | Supporting transfusion decisions

Beyond prediction of the likelihood of transfusion, ML can identify inappropriate transfusions, recognize patient groups by predicted transfusion outcomes, and enable precise dosing of blood products in efforts to reduce iron overload.[71,72] Through the analysis of existing clinical trials data, ML enabled estimates of the causal effect of preoperative plasma transfusion on perioperative bleeding in patients with a high International Normalized Ratio test result[73] and of different ratios of platelets and plasma relative to PRBC on mortality and hemostasis in trauma patients.[74] Bruun-Rasmussen et al.[75] used ML to emulate a randomized controlled trial in the context of sex-matched transfusion policy. Ngufor et al.[76] take a key step toward personalized medicine, clustering patients using unsupervised ML to determine whether they will benefit from plasma transfusion. Models to identify inappropriate transfusions may reduce the labor required for retrospective quality control[77] and support local efforts to reduce unnecessary orders and transfusions.[78] It may also be possible to identify situations where ongoing transfusion is futile, but this has proved challenging.[79]

TABLE 2 Summary of methodological considerations by category.

Machine learning methods												
Transfusion topic	**Subgroup**	***N***	**Multisite (%)**	**Source data EHR (%)**	**Tree-based (%)**	**NN (%)**	**Other (%)**	**Prospective evaluation (%)**	**Outcome performance comparator (%)**	**Reporting framework stated (%)**	**Data available (%)**	**Code available (%)**
Predicting transfusion	GI bleeding	2	2 (100)	2 (100)	1 (50)	2 (100)	2 (100)	0 (0)	2 (100)	1 (50)	2 (100)	2 (100)
	Hemato/oncology	1	0 (0)	1 (100)	0 (0)	1 (100)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	Obstetrics	2	0 (0)	2 (100)	2 (100)	0 (0)	1 (50)	0 (0)	2 (100)	0 (0)	1 (50)	0 (0)
	Other hospital settings	3	1 (33)	3 (100)	3 (100)	2 (67)	2 (67)	0 (0)	2 (67)	0 (0)	2 (67)	0 (0)
	Surgery (Orthopedic)	8	5 (63)	7 (88)	8 (100)	3 (38)	7 (88)	1 (13)	6 (75)	1 (13)	4 (50)	1 (13)
	Surgery (Cardiac)	5	1 (20)	4 (80)	4 (80)	2 (40)	4 (80)	0 (0)	2 (40)	2 (40)	1 (20)	0 (0)
	Surgery (Spinal)	4	2 (50)	2 (50)	3 (75)	2 (50)	1 (25)	0 (0)	1 (25)	1 (25)	3 (75)	0 (0)
	Surgery (Liver transplant)	3	2 (67)	3 (100)	3 (100)	2 (67)	2 (67)	2 (67)	2 (67)	1 (33)	1 (33)	0 (0)
	Surgery (Multiple specialties)	7	4 (57)	2 (29)	4 (57)	3 (43)	4 (57)	0 (0)	5 (71)	2 (29)	2 (29)	2 (29)
	Surgery (Other)	6	2 (33)	5 (83)	3 (50)	5 (83)	4 (67)	0 (0)	5 (83)	1 (17)	3 (50)	1 (17)
	Trauma	13	5 (38)	10 (77)	7 (54)	4 (31)	10 (77)	2 (15)	9 (69)	2 (15)	2 (15)	0 (0)
	All	54	24 (44)	41 (76)	38 (70)	26 (48)	37 (69)	5 (9)	36 (67)	11 (20)	21 (39)	6 (11)
Supporting transfusion decisions	Adverse outcome	4	2 (50)	3 (75)	4 (100)	1 (25)	4 (100)	0 (0)	2 (50)	1 (25)	1 (25)	1 (25)
	Appropriateness	3	2 (67)	3 (100)	2 (67)	1 (33)	1 (33)	1 (33)	1 (33)	0 (0)	1 (33)	1 (33)
	Dose	2	1 (50)	0 (0)	1 (50)	1 (50)	2 (100)	0 (0)	1 (50)	0 (0)	1 (50)	1 (50)
	All	9	5 (56)	6 (67)	7 (78)	3 (33)	7 (78)	1 (11)	4 (44)	1 (11)	3 (33)	3 (33)
Transfusion safety	Hemovigilance	13	4 (31)	11 (85)	11 (85)	4 (31)	9 (69)	0 (0)	9 (69)	0 (0)	1 (8)	1 (8)
	Laboratory	8	1 (13)	0 (0)	3 (38)	5 (63)	3 (38)	1 (13)	5 (63)	N/A	1 (13)	1 (13)
	All	21	5 (24)	11 (52)	14 (67)	9 (43)	12 (57)	1 (5)	14 (67)	N/A	2 (10)	2 (10)
Hospital blood bank	Ordering	7	2 (29)	5 (71)	3 (43)	2 (29)	5 (71)	1 (14)	6 (86)	N/A	0 (0)	0 (0)
	Waste	2	2 (100)	2 (100)	1 (50)	0 (0)	2 (100)	0 (0)	0 (0)	N/A	0 (0)	0 (0)

TABLE 2 (Continued)

Machine learning methods												
Transfusion topic	Subgroup	*N*	Multisite (%)	Source data EHR (%)	Tree-based (%)	NN (%)	Other (%)	Prospective evaluation (%)	Outcome performance comparator (%)	Reporting framework stated (%)	Data available (%)	Code available (%)
	All	9	4 (44)	7 (78)	4 (44)	2 (22)	7 (78)	1 (11)	6 (67)	N/A	0 (0)	0 (0)
Total		93	38 (41)	65 (70)	63 (68)	40 (43)	63 (68)	8 (9)	60 (65)	N/A	26 (28)	11 (12)

Abbreviations: EHR, electronic health record; GI, gastrointestinal; NN, neural network.

3.8 | Transfusion safety

Identified studies in this category were divided into hemovigilance and laboratory support in the Blood Bank.

3.9 | Hemovigilance

ML has been applied primarily to enhance the ability to detect and predict acute transfusion reactions (ATRs) and adverse transfusion events. Novel information retrieval methods, such as natural language processing (NLP), when applied to electronic health records (EHRs) have demonstrated underreporting by clinicians and the potential to improve detection.[80–82] Alternatively, Roubinian et al.[83] and Nguyen et al.[84] incorporated novel biomarkers into classification models and decision tree analysis, respectively.

While the focus of ML is on prediction, and a causal relationship cannot be assumed of the covariates found to have high predictive value, identification of novel risk factors for hypothesis generation and further research can be useful as seen in transfusion-associated lung injury (TRALI)[85] and in pediatric transfusion-associated hyperkalemia.[86] Recognizing that transparency and accountability are essential for clinicians in generating hypotheses, Zhu et al.[87,88] focus on explainable AI when presenting adverse events during neonatal hyperbilirubinemia exchange transfusion, particularly through use of SHapley Additive exPlanation (SHAP).

3.9.1 | Blood group identification and other laboratory applications

In a laboratory setting, 63% (six of eight) studies investigated the use of ML to assist blood group identification including two to help classify antibody reactions where ambiguity in human interpretation exists.[89,90] Doan et al.[91] and Kim et al.[92] introduced image-based deep learning as a novel approach to perform phenotype assessment of red blood cell storage lesions to predict red cell quality prior to transfusion.

3.10 | Hospital blood bank inventory management

The availability of EHRs has supported the development of models to forecast blood product demand and recommend order quantities, based on aggregated patient data in addition to historical demand patterns.[93–99] A model implemented in a Canadian hospital for PRBC reduced

wastage and daily stockholding,[93] and it is common for forecasting models to be implemented into simulations to estimate the potential benefits of deploying them.[94–97] In addition to studies supporting ordering, two studies investigated the use of ML to directly address wastage in a hospital blood bank by predicting discards[100] and identifying transaction patterns associated with wastage.[101]

3.11 | Methodological considerations

Selected characteristics of the methodology of identified studies are summarized in Table 2.

Only 41% of all studies were multisite. Source data were derived from EHRs in 70%, defined as hospital data that were collected routinely without research or audit intent. Alternative data sources included research databases and laboratory primary research data.

All but two papers (98%) used supervised learning, while only four papers (4%) used unsupervised learning methods. None of the papers used reinforcement learning. In Table 2, we divide supervised ML methods into three broad groups: tree-based methods, neural networks, and other methods. We describe these groups in the Appendix. If a study used an ensemble of techniques from more than one of these categories, then the underlying techniques are counted as having been used in that study. Tree-based models were included in 68% of the studies and neural networks in 43% of the studies.

A common approach taken by many of the studies is to compare several different ML methods readily available in software libraries such as Python's scikit-learn. A small number of papers investigate novel ML ideas including the use of a secondary model to provide a confidence level in predictions[57] and using weak labels in the form of age information to train a model to classify red cell quality without relying on subjective human expert labels.[91]

Only eight studies (9%) reported the results of prospective evaluation or deployment. Of these, four evaluated their models on prospectively collected data,[35,37,90,102] one conducted a "shadow test" in which predictions were generated in real time for evaluation but not used for decision-making,[56] and three describe implementation as part of live decision-making.[65,78,93] A majority of studies (65%) included an outcome performance comparator, defined as a logistic or linear regression model, a previously reported method for the same problem, or a baseline representing current practice.

Where expected, as with individual-level prediction, reference to a reporting framework was infrequent. Recognized reporting frameworks including TRIPOD and STROBE were utilized in only 11/54 studies within predicting transfusion. None of the work predicts transfusion reactions or adverse events within the transfusion safety subgroup of hemovigilance reported in accordance with a recognized framework. Data were stated to be available in 28% of studies and code in only 12%, which will limit future researchers' ability to reproduce and extend the work performed to date. Yamada et al.[86] provided their full data analytic protocol as an electronic notebook, an example supporting reproducibility and open science.

4 | DISCUSSION

Research in the field of ML in transfusion is expanding rapidly with exciting applications as evidenced by the number of publications. However, our review also highlights clear challenges surrounding transparency, interpretability, and generalizability of findings. Most studies are single center and have no prospective validation or implementation. ML model code and data are rarely made available for external validation, and there is limited justification of methods, with best performing models often selected from a trial of those commonly available.

Where ML performance characteristics are often encouraging, the authors emphasize caution in interpretation of evidence that models can achieve improved performance as compared to current practice. As within predicting transfusion, ML does not always offer advantage over LR when task-specific performance is compared and demonstrates the difficultly in interpreting the clinical potential of ML while tasks, reporting measures, and methodology remain so variable. While challenging, prospective deployment of a model within the clinical workflow and subsequent evaluation of changes in key performance indicators is highly desirable in ML. Our findings of limited prospective testing and deployment are consistent with those of the wider field where translation remains a challenge, and researchers are producing frameworks[103] and sharing case studies[104] to help close this gap. A simulated workflow, as developed in recent studies,[12,105] as a method for evaluating the potential impact of a model may help to prioritize candidates for prospective testing.

Data were extracted from EHRs or from medical devices in three of the four studies where predictions were made in real time, as part of a live workflow or a "shadow test,"[56,78,93] while the remaining case required data to be entered manually into a smartphone application.[65] The latter may face fewer initial barriers to deployment such challenges involved in integrating different systems, but there is a risk of manual data entry errors

and a limit on the quantity and types of data that can be quickly and accurately entered. Recent developments in NLP, such as ChatGPT, may lead to the development of systems that can interactively support decision-making and integrate structured data from EHRs with clinical notes, and brief-written (or transcribed verbal) descriptions of the problem at hand.[106]

Current variation in transfusion practice, particularly in prediction where the outcome "decision to transfuse" remains clinician-dependent, could perpetuate suboptimal practices. Models may embed class imbalance and generate predictions enriched by patient episodes, which are elsewhere considered to be over-transfused. This problem may be reflected in poor external validation of pretrained models if the sites use different guidance or practice.[68] Although considered beyond the scope of this review, it may be of interest to review studies where the contribution of physicians, for example, surgeons and anesthesiologists, features as a variable of the model, or where variables behind physicians' decision-making are explored in more detail,[44] to address reasons for variation of practice. Such an approach could prompt action to address discrepancies, particularly as the large, multicenter datasets used for ML are also well suited to address physician effects while preserving anonymity. Clear key performance indicators remain an unmet need in transfusion by which to evaluate clinically relevant outcomes following transfusion in a standardized way.

While practice variation impacts the generalizability of trained models, underlying methods may still generalize to new sites if retrained on local data. Additionally, successful integration of a model into the workflow may change patterns that have been learned (e.g., which tests are ordered and how often). It is crucial to continuously monitor the predictions of a deployed model to ensure that its predictions remain valid and useful.[107] The ability to fine-tune and update models using local[108] and/or more recent data[109] offers huge potential advantage of ML-based predictive modeling over historic simplified scores and static prediction rules in addressing these challenges, enabling models to accommodate new clinical trends and evaluate performance as compared to current practice in an iterative manner.[23]

To our knowledge, this review is the first attempt to collate the literature on a wide range of applications of ML in transfusion medicine. Our analysis extends the work of Meier and Tschoellitsch[13] who describe 47 articles of ML applied to PBM (including bleeding and anemia) from a 2021 PubMed search. We have captured information on emerging areas of interest to clinicians and researchers, and by review of ongoing challenges faced in the interpretation and translation of ML, we also offer suggested priorities for future reporting and work.

Our study has a number of limitations. The heterogeneity of methods and infrequent use of reporting frameworks makes synthesis of results and interpretation challenging, as well as creating barriers for researchers to build upon and validate outcomes. Researchers should be encouraged to provide work as an online open-source repository and share computational tools.[30,110] Developing common task definitions and following established reporting frameworks would make it easier to compare methods and identify candidates for prospective validation and subsequent implementation. Secondly, in setting out to give an overview of the literature, the volume of publications captured while maintaining broad search terms meant it was beyond the scope of this review to extend to multiple databases and we acknowledge that relevant studies may have been missed. Citation review was performed in efforts to minimize this. Further studies might benefit from more focused reviews on selected themes in transfusion medicine. Lastly, while we apply the "main outcome of transfusion" to identify studies to support PBM, we recognize that the concept of PBM goes well beyond this such as optimization of anemia (e.g., erythropoietin therapy and iron therapy) and that these areas deserve exploration in future studies. As the body of literature of ML in transfusion and PBM grows, so will the potential for more focused systematic reviews. This review adds to the continuously evolving, contemporaneous studies and reviews essential to engage clinicians new to the idea of ML.[15,111]

5 | CONCLUSION

There has been a major expansion of the literature in recent years, reflecting the interest and enthusiasm toward the application of ML in transfusion medicine. However, many challenges and limitations remain to include data quality and access, adherence to (and existence of) appropriate reporting frameworks, and generalizability of findings. Emphasis should be on consistent reporting, sharing of code, and prospective validation with comparison to current practice of future studies.

FUNDING INFORMATION

This study was supported by the NIHR Blood and Transplant Research Unit in Data Driven Transfusion Practice (NIHR203334), UKRI Training Grant No. EP/S021612/1, the Centre for Doctoral Training in AI-enabled Healthcare Systems, and the NIHR University College London Hospitals Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

CONFLICT OF INTEREST STATEMENT

The authors have disclosed no conflicts of interest.

ORCID

Suzanne Maynard https://orcid.org/0000-0003-3039-2562

REFERENCES

1. Tinegate H, Pendry K, Murphy M, Babra P, Grant-Casey J, Hopkinson C, et al. Where do all the red blood cells (RBCs) go? Results of a survey of RBC use in England and North Wales in 2014. Transfusion. 2016;56(1):139–45. https://doi.org/10.1111/trf.13342
2. European Committee on Blood Transfusion. The collection, testing and use of blood and blood components in Europe. 2017, 2018 and 2019 report. Cited 2023 July 24. Available from: https://www.edqm.eu/en/reports-blood
3. Delaney M, Wendel S, Bercovitz RS, Cid J, Cohn C, Dunbar NM, et al. Transfusion reactions: prevention, diagnosis, and treatment. Lancet. 2016;388(10061):2825–36. https://doi.org/10.1016/S0140-6736(15)01313-6
4. Desborough MJR, Hockley B, Sekhar M, Burroughs AK, Stanworth SJ, Jairath V, et al. Patterns of blood component use in cirrhosis: a nationwide study. Liver Int. 2016;36(4):522–9. https://doi.org/10.1111/liv.12999
5. Bruun MT, Pendry K, Georgsen J, Manzini P, Lorenzi M, Wikman A, et al. Patient Blood Management in Europe: surveys on top indications for red blood cell use and Patient Blood Management organization and activities in seven European university hospitals. Vox Sang. 2016;111(4):391–8. https://doi.org/10.1111/vox.12435
6. Stanworth SJ, New HV, Apelseth TO, Brunskill S, Cardigan R, Doree C, et al. Effects of the COVID-19 pandemic on supply and use of blood for transfusion. Lancet Haematol. 2020;7(10):e756–64. https://doi.org/10.1016/S2352-3026(20)30186-1
7. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA. 2018;319(13):1317–8. https://doi.org/10.1001/jama.2017.18391
8. Nwanosike EM, Conway BR, Merchant HA, Hasan SS. Potential applications and performance of machine learning techniques and algorithms in clinical practice: a systematic review. Int J Med Inform. 2022;159:104679. https://doi.org/10.1016/j.ijmedinf.2021.104679
9. Vercio LL, Amador K, Bannister JJ, Crites S, Gutierrez A, MacDonald ME, et al. Supervised machine learning tools: a tutorial for clinicians. J Neural Eng. 2020;17(6):062001. https://doi.org/10.1088/1741-2552/abbff2
10. Al-Zaiti SS, Alghwiri AA, Hu X, Clermont G, Peace A, Macfarlane P, et al. A clinician's guide to understanding and critically appraising machine learning studies: a checklist for Ruling Out Bias Using Standard Tools in Machine Learning (ROBUST-ML). Eur Heart J Digit Health. 2022;3(2):125–40. https://doi.org/10.1093/ehjdh/ztac016
11. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. Lancet Oncol. 2019;20(5):e262–73. https://doi.org/10.1016/S1470-2045(19)30149-4
12. Wornow M, Gyang Ross E, Callahan A, Shah NH. APLUS: a Python library for usefulness simulations of machine learning models in healthcare. J Biomed Inform. 2023;139:104319. https://doi.org/10.1016/j.jbi.2023.104319
13. Meier JM, Tschoellitsch T. Artificial intelligence and machine learning in patient blood management: a scoping review. Anesth Analg. 2022;135(3):524–31. https://doi.org/10.1213/ANE.0000000000006047
14. Government of Canada CI of HR. A guide to knowledge synthesis – CIHR. Published March 25, 2010 [cited 2023 July 24]. Available from: https://cihr-irsc.gc.ca/e/41382.html
15. Ben-Israel D, Jacobs WB, Casha S, Lang S, Ryu WHA, de Lotbiniere-Bassett M, et al. The impact of machine learning on patient care: a systematic review. Artif Intell Med. 2020;103:101785. https://doi.org/10.1016/j.artmed.2019.101785
16. Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. A review of challenges and opportunities in machine learning for health. AMIA Jt Summits Transl Sci Proc. 2020;2020:191–200.
17. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med. 2019;380(14):1347–58. https://doi.org/10.1056/NEJMra1814259
18. Rowe M. An introduction to machine learning for clinicians. Acad Med. 2019;94(10):1433–6. https://doi.org/10.1097/ACM.0000000000002792
19. Chang CM, Hung JH, Hu YH, Lee PJ, Shen CC. Prediction of preoperative blood preparation for orthopedic surgery patients: a supervised learning approach. Appl Sci. 2018;8(9):1559. https://doi.org/10.3390/app8091559
20. Cohen-Levy WB, Klemt C, Tirumala V, Burns JC, Barghi A, Habibi Y, et al. Artificial neural networks for the prediction of transfusion rates in primary total hip arthroplasty. Arch Orthop Trauma Surg. 2023;143(3):1643–50. https://doi.org/10.1007/s00402-022-04391-8
21. Gowd AK, Agarwalla A, Amin NH, Romeo AA, Nicholson GP, Verma NN, et al. Construct validation of machine learning in the prediction of short-term postoperative complications following total shoulder arthroplasty. J Shoulder Elbow Surg. 2019;28(12):e410–21. https://doi.org/10.1016/j.jse.2019.05.017
22. Huang Z, Huang C, Xie J, Ma J, Cao GR, Huang Q, et al. Analysis of a large data set to identify predictors of blood transfusion in primary total hip and knee arthroplasty. Transfusion. 2018;58(8):1855–62. https://doi.org/10.1111/trf.14783
23. Huang Z, Martin J, Huang Q, Ma J, Pei F, Huang C. Predicting postoperative transfusion in elective total HIP and knee arthroplasty: comparison of different machine learning models of a case-control study. Int J Surg. 2021;96:106183. https://doi.org/10.1016/j.ijsu.2021.106183
24. Jo C, Ko S, Shin WC, Han HS, Lee MC, Ko T, et al. Transfusion after total knee arthroplasty can be predicted using the machine learning algorithm. Knee Surg Sports Traumatol Arthrosc. 2019;28(6):1757–64. https://doi.org/10.1007/s00167-019-05602-3
25. Mohammed H, Huang Y, Memtsoudis S, Parks M, Huang Y, Ma Y. Utilization of machine learning methods for predicting surgical outcomes after total knee arthroplasty. PloS One. 2022;17(3):e0263897. https://doi.org/10.1371/journal.pone.0263897
26. Cevenini G, Barbini E, Massai MR, Barbini P. A naïve Bayes classifier for planning transfusion requirements in heart

surgery. J Eval Clin Pract. 2013;19(1):25–9. https://doi.org/10.1111/j.1365-2753.2011.01762.x
27. Liu S, Zhou R, Xia XQ, Ren H, Wang LY, Sang RR, et al. Machine learning models to predict red blood cell transfusion in patients undergoing mitral valve surgery. Ann Transl Med. 2021;9(7):530. https://doi.org/10.21037/atm-20-7375
28. Tschoellitsch T, Böck C, Maletić TT, Hofmann A, Meier J. Machine learning-based prediction of massive perioperative allogeneic blood transfusion in cardiac surgery. Eur J Anaesthesiol. 2022;39(9):766–73. https://doi.org/10.1097/EJA.0000000000001721
29. Wang Z, Zhe S, Zimmerman J, Morrisey C, Tonna JE, Sharma V, et al. Development and validation of a machine learning method to predict intraoperative red blood cell transfusions in cardiothoracic surgery. Sci Rep. 2022;12(1):1355. https://doi.org/10.1038/s41598-022-05445-y
30. Gurm HS, Kooiman J, LaLonde T, Grines C, Share D, Seth M. A random forest based risk model for reliable and accurate prediction of receipt of transfusion in patients undergoing percutaneous coronary intervention. PloS One. 2014;9(5):e96385. https://doi.org/10.1371/journal.pone.0096385
31. De la Garza RR, Hamad MK, Ryvlin J, Krol O, Passias PG, Fourman MS, et al. An artificial neural network model for the prediction of perioperative blood transfusion in adult spinal deformity surgery. J Clin Med. 2022;11(15):4436. https://doi.org/10.3390/jcm11154436
32. Dong S, Li W, Tang ZR, Wang H, Pei H, Yuan B. Development and validation of a novel predictive model and web calculator for evaluating transfusion risk after spinal fusion for spinal tuberculosis: a retrospective cohort study. BMC Musculoskelet Disord. 2021;22(1):825. https://doi.org/10.1186/s12891-021-04715-6
33. Durand WM, Depasse JM, Daniels AH. Predictive modeling for blood transfusion after adult spinal deformity surgery. Spine. 2018;43(15):1058–66. https://doi.org/10.1097/BRS.0000000000002515
34. Raman T, Vasquez-Montes D, Varlotta C, Passias PG, Errico TJ. Decision tree-based modelling for identification of predictors of blood loss and transfusion requirement after adult spinal deformity surgery. Int J Spine Surg. 2020;14(1):87–95. https://doi.org/10.14444/7012
35. Chen S, Liu L-P, Wang Y-J, Zhou X-H, Dong H, Chen Z-W, et al. Advancing prediction of risk of intraoperative massive blood transfusion in liver transplantation with machine learning models. A multicenter retrospective study. Front Neuroinform. 2022;16:893452. https://doi.org/10.3389/fninf.2022.893452
36. Cywinski JB, Alster JM, Miller C, Vogt DP, Parker BM. Prediction of intraoperative transfusion requirements during orthotopic liver transplantation and the influence on postoperative patient survival. Anesth Analg. 2014;118(2):428–37. https://doi.org/10.1213/ANE.0b013e3182a76f19
37. Liu LP, Zhao QY, Wu J, Luo YW, Dong H, Chen ZW, et al. Machine learning for the prediction of red blood cell transfusion in patients during or after liver transplantation surgery. Front Med. 2021;8:632210. https://doi.org/10.3389/fmed.2021.632210
38. Feng Y, Xu Z, Sun X, Wang D, Yu Y. Machine learning for predicting preoperative red blood cell demand. Transfus Med. 2021;31(4):262–70. https://doi.org/10.1111/tme.12794
39. Hayn D, Kreiner K, Kastner P, Breznik N, Hofmann A, Gombotz H, et al. Data driven methods for predicting blood transfusion needs in elective surgery. In: *Health Informatics Meets eHealth: Predictive Modeling in Healthcare – From Prediction to Prevention. Proceedings of the 10th eHealth Conference 2016*. IOS Press; 2016. p. 9–16.
40. Hayn D, Kreiner K, Ebner H, Kastner P, Breznik N, Rzepka A, et al. Development of multivariable models to predict and benchmark transfusion in elective surgery supporting patient blood management. Appl Clin Inform. 2017;8(2):617–31. https://doi.org/10.4338/ACI-2016-11-RA-0195
41. Lou SS, Liu H, Lu C, Wildes TS, Hall BL, Kannampallil T. Personalized surgical transfusion risk prediction using machine learning to guide preoperative type and screen orders. Anesthesiology. 2022;137(1):55–66. https://doi.org/10.1097/ALN.0000000000004139
42. Ngufor C, Upadhyaya S, Murphree D, Kor D, Pathak J. Multi-task learning with selective cross-task transfer for predicting bleeding and other important patient outcomes. In: *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. 2015. p. 1–8. https://doi.org/10.1109/DSAA.2015.7344836
43. Walczak S, Velanovich V. Prediction of perioperative transfusions using an artificial neural network. PloS One. 2020;15(2):e0229450. https://doi.org/10.1371/journal.pone.0229450
44. Etchells TA, Harrison MJ. Orthogonal search-based rule extraction for modelling the decision to transfuse. Anaesthesia. 2006;61(4):335–8. https://doi.org/10.1111/j.1365-2044.2006.04545.x
45. Jalali A, Lonsdale H, Zamora LV, Ahumada L, Nguyen ATH, Rehman M, et al. Machine learning applied to registry data: development of a patient-specific prediction model for blood transfusion requirements during craniofacial surgery using the pediatric craniofacial perioperative registry dataset. Anesth Analg. 2021;132(1):160–71. https://doi.org/10.1213/ANE.0000000000004988
46. Tunthanathip T, Sae-heng S, Oearsakul T, Kaewborisutsakul A, Taweesomboonyat C. Economic impact of a machine learning-based strategy for preparation of blood products in brain tumor surgery. PloS One. 2022;17(7):e0270916. https://doi.org/10.1371/journal.pone.0270916
47. Dhiman P, Ma J, Gibbs VN, Rampotas A, Kamal H, Arshad SS, et al. Systematic review highlights high risk of bias of clinical prediction models for blood transfusion in patients undergoing elective surgery. J Clin Epidemiol. 2023;159:10–30. https://doi.org/10.1016/j.jclinepi.2023.05.002
48. El-Menyar A, Mekkodathil A, Abdelrahman H, Latifi R, Galwankar S, Al-Thani H, et al. Review of existing scoring systems for massive blood transfusion in trauma patients: where do we stand? Shock. 2019;52(3):288–99. https://doi.org/10.1097/SHK.0000000000001359
49. Maegele M, Brockamp T, Nienaber U, Probst C, Schoechl H, Goerlinger K, et al. Predictive models and algorithms for the need of transfusion including massive transfusion in severely injured patients. Transfus Med Hemother. 2012;39(2):85–97. https://doi.org/10.1159/000337243
50. Tran A, Matar M, Lampron J, Steyerberg E, Taljaard M, Vaillancourt C. Early identification of patients requiring massive transfusion, embolization or hemostatic surgery for traumatic hemorrhage: a systematic review and meta-analysis. J Trauma Acute Care Surg. 2018;84(3):505–16. https://doi.org/10.1097/TA.0000000000001760

51. Yin G, Radulovic N, O'Neill M, Lightfoot D, Nolan B. Predictors of transfusion in trauma and their utility in the prehospital environment: a scoping review. Prehosp Emerg Care. 2022; 27(5):575–85. https://doi.org/10.1080/10903127.2022.2120935
52. Committee on Trauma of the American College of Surgeons. ACS TQIP Massive Transfusion in Trauma Guidelines. 2014. [cited 2023 April 26]. Available from: https://www.facs.org/media/zcjdtrd1/transfusion_guildelines.pdf
53. Meyer DE, Vincent LA, Fox EE, O'Keeffe T, Inaba K, Bulger E, et al. Every minute counts: time to delivery of initial massive transfusion cooler and its impact on mortality. J Trauma Acute Care Surg. 2017;83(1):19–24. https://doi.org/10.1097/TA.0000000000001531
54. Mina MJ, Winkler AM, Dente CJ. Let technology do the work: improving prediction of massive transfusion with the aid of a smartphone application. J Trauma Acute Care Surg. 2013; 75(4):669–75. https://doi.org/10.1097/TA.0b013e3182a12ba6
55. Avital G, Gelikas S, Radomislensky I, Tsur AM, Sorkin A, Shinar E, et al. A prehospital scoring system for predicting the need for emergent blood product transfusion. Transfusion. 2021;61(S1):S195–205. https://doi.org/10.1111/trf.16529
56. Liu J, Khitrov MY, Gates JD, Odom SR, Havens JM, de Moya MA, et al. Automated analysis of vital signs to identify patients with substantial bleeding before hospital arrival: a feasibility study. Shock. 2015;43(5):429–36. https://doi.org/10.1097/SHK.0000000000000328
57. Nederpelt CJ, Mokhtari AK, Alser O, Tsiligkaridis T, Roberts J, Cha M, et al. Development of a field artificial intelligence triage tool: confidence in the prediction of shock, transfusion, and definitive surgical therapy in patients with truncal gunshot wounds. J Trauma Acute Care Surg. 2021; 90(6):1054–60. https://doi.org/10.1097/TA.0000000000003155
58. Christie SA, Conroy AS, Callcut RA, Hubbard AE, Cohen MJ. Dynamic multi-outcome prediction after injury: applying adaptive machine learning for precision medicine in trauma. PloS One. 2019;14(4):e0213836. https://doi.org/10.1371/journal.pone.0213836
59. Feng YN, Xu ZH, Liu JT, Sun XL, Wang DQ, Yu Y. Intelligent prediction of RBC demand in trauma patients using decision tree methods. Mil Med Res. 2021;8(1):33. https://doi.org/10.1186/s40779-021-00326-3
60. Walczak S. Artificial neural network medical decision support tool: predicting transfusion requirements of ER patients. IEEE Trans Inf Technol Biomed. 2005;9(3):468–74. https://doi.org/10.1109/TITB.2005.847510
61. Seheult JN, Anto VP, Farhat N, Stram MN, Spinella PC, Alarcon L, et al. Application of a recursive partitioning decision tree algorithm for the prediction of massive transfusion in civilian trauma: the MTPitt prediction tool. Transfusion. 2019;59(3):953–64. https://doi.org/10.1111/trf.15078
62. Shahi N, Shahi AK, Phillips R, Shirek G, Bensard D, Moulton SL. Decision-making in pediatric blunt solid organ injury: a deep learning approach to predict massive transfusion, need for operative management, and mortality risk. J Pediatr Surg. 2021;56(2):379–84. https://doi.org/10.1016/j.jpedsurg.2020.10.021
63. Liu W, Wu C, Zou J, Le A. Massive transfusion prediction in patients with multiple trauma by decision tree: a retrospective analysis. Indian J Hematol Blood Transfus. 2021;37(2):302–8. https://doi.org/10.1007/s12288-020-01348-y
64. Hodgman EI, Cripps MW, Mina MJ, Bulger EM, Schreiber MA, Brasel KJ, et al. External validation of a smartphone app model to predict the need for massive transfusion using five different definitions. J Trauma Acute Care Surg. 2018;84(2):397–402. https://doi.org/10.1097/TA.0000000000001756
65. Dente CJ, Mina MJ, Morse BC, Hensman H, Schobel S, Gelbard RB, et al. Predicting the need for massive transfusion: prospective validation of a smartphone-based clinical decision support tool. Surgery. 2021;170(5):1574–80. https://doi.org/10.1016/j.surg.2021.04.034
66. Say L, Chou D, Gemmill A, Tunçalp Ö, Moller AB, Daniels J, et al. Global causes of maternal death: a WHO systematic analysis. Lancet Glob Health. 2014;2(6):e323–33. https://doi.org/10.1016/S2214-109X(14)70227-X
67. Levi R, Carli F, Arévalo AR, Altinel Y, Stein DJ, Naldini MM, et al. Artificial intelligence-based prediction of transfusion in the intensive care unit in patients with gastrointestinal bleeding. BMJ Health Care Inform. 2021;28:e100245.
68. Shung D, Huang J, Castro E, Tay JK, Simonov M, Laine L, et al. Neural network predicts need for red blood cell transfusion for patients with acute gastrointestinal bleeding admitted to the intensive care unit. Sci Rep. 2021;11:8827. https://doi.org/10.1038/s41598-021-88226-3
69. Lee S, Lee S, Chang D, Song MH, Kim JY, Lee S. Explainable machine learning based a packed red blood cell transfusion prediction and evaluation for major internal medical condition. J Inf Process Syst. 2022;18(3):302–10.
70. Ghassemi M, Wu M, Hughes MC, Szolovits P, Doshi-Velez F. Predicting intervention onset in the ICU with switching state space models. In: *AMIA Joint Summits on Translational Science Proceedings*. 2017. p. 82–91.
71. Epah J, Gülec I, Winter S, Dörr J, Geisen C, Haecker E, et al. From unit to dose: a machine learning approach for precise prediction of hemoglobin and iron content in individual packed red blood cell units. Adv Sci. 2022;9(36):2204077. https://doi.org/10.1002/advs.202204077
72. Zhang M, Fu Z, Hou X, Li G, Lin L. Improving the quantitative analysis accuracy of bagged liquid components with strong scattering by multi-pathlength data fusion. Infrared Phys Technol. 2019;99:39–44. https://doi.org/10.1016/j.infrared.2019.04.006
73. Ngufor C, Murphree D, Upadhyaya S, Madde N, Kor D, Pathak J. Effects of plasma transfusion on perioperative bleeding complications: a machine learning approach. Stud Health Technol Inform. 2015;216:721–5.
74. Nguyen M, Pirracchio R, Kornblith LZ, Callcut R, Fox EE, Wade CE, et al. Dynamic impact of transfusion ratios on outcomes in severely injured patients: targeted machine learning analysis of the Pragmatic, Randomized Optimal Platelet and Plasma Ratios randomized clinical trial. J Trauma Acute Care Surg. 2020;89(3):505–13. https://doi.org/10.1097/TA.0000000000002819
75. Bruun-Rasmussen P, Andersen PK, Banasik K, Brunak S, Johansson PI. Estimating the effect of donor sex on red blood cell transfused patient mortality: a retrospective cohort study using a targeted learning and emulated trials-based approach. eClinicalMedicine. 2022;51:101628. https://doi.org/10.1016/j.eclinm.2022.101628

76. Ngufor C, Warner MA, Murphree DH, Liu H, Carter R, Storlie CB, et al. Identification of clinically meaningful plasma transfusion subgroups using unsupervised random forest clustering. AMIA Annu Symp Proc. 2018;2017:1332–41.
77. Yao Y, Cifuentes J, Zheng B, Yan M. Computer algorithm can match physicians' decisions about blood transfusions. J Transl Med. 2019;17(1):1–5. https://doi.org/10.1186/s12967-019-2085-y
78. McGlothlin JP, Crawford E, Srinivasan H, Cianci C, Bruneau B, Dorotta I. Reducing red blood cell transfusions. In: Bursa M, Holzinger A, Renda ME, Khuri S, editors. Information technology in bio- and medical informatics. Lecture notes in computer science. Cham: Springer International Publishing; 2017. p. 120–33. https://doi.org/10.1007/978-3-319-64265-9_10
79. Barbosa RR, Rowell SE, Diggs BS, Schreiber MA, Trauma Outcomes Group, Holcomb JB, et al. Profoundly abnormal initial physiologic and biochemical data cannot be used to determine futility in massively transfused trauma patients. J Trauma Acute Care Surg. 2011;71(2):S364–9. https://doi.org/10.1097/TA.0b013e318227f170
80. Tsatsoulis C. Finding clusters of similar events within clinical incident reports: a novel methodology combining case based reasoning and information retrieval. Qual Saf Health Care. 2003;12(90002):24ii–2432ii. https://doi.org/10.1136/qhc.12.suppl_2.ii24
81. Whitaker B, Pizarro J, Deady M, Williams A, Ezzeldin H, Belov A, et al. Detection of allergic transfusion-related adverse events from electronic medical records. Transfusion. 2022;62(10):2029–38. https://doi.org/10.1111/trf.17069
82. Clifford L, Singh A, Wilson GA, Toy P, Gajic O, Malinchoc M, et al. Electronic health record surveillance algorithms facilitate the detection of transfusion-related pulmonary complications. Transfusion. 2013;53(6):1205–16. https://doi.org/10.1111/j.1537-2995.2012.03886.x
83. Roubinian NH, Chowdhury D, Hendrickson JE, Triulzi DJ, Gottschall JL, Looney MR, et al. NT-proBNP levels in the identification and classification of pulmonary transfusion reactions. Transfusion. 2020;60(11):2548–56. https://doi.org/10.1111/trf.16059
84. Nguyen KA, Hamzeh-Cognasse H, Sebban M, Fromont E, Chavarin P, Absi L, et al. A computerized prediction model of hazardous inflammatory platelet transfusion outcomes. PloS One. 2014;9(5):e97082. https://doi.org/10.1371/journal.pone.0097082
85. Torres MD, Torres A, Cuellar F, de la Luz Torres M, de León EP, Pinales F. Evolutionary computation in the identification of risk factors. Case of TRALI. Expert Syst Appl. 2014;41(3):831–40. https://doi.org/10.1016/j.eswa.2013.08.013
86. Yamada C, Edelson MF, Lee AC, Saifee NH, Dinov ID. Transfusion-associated hyperkalemia in pediatric population: analyses for risk factors and recommendations. Transfusion. 2022;62(12):2503–14. https://doi.org/10.1111/trf.17135
87. Zhu S, Zhou L, Feng Y, Zhu J, Shu Q, Li H. Understanding the risk factors for adverse events during exchange transfusion in neonatal hyperbilirubinemia using explainable artificial intelligence. BMC Pediatr. 2022;22(1):567. https://doi.org/10.1186/s12887-022-03615-5
88. Huang AA, Huang SY. Increasing transparency in machine learning through bootstrap simulation and shapely additive explanations. PloS One. 2023;18(2):e0281922. https://doi.org/10.1371/journal.pone.0281922
89. Wu K, Duan S, Wang Y, Wang H, Gao X. Convolutional neural network-based automatic classification for incomplete antibody reaction intensity in solid phase anti-human globulin test image. Med Biol Eng Comput. 2022;60(4):1211–22. https://doi.org/10.1007/s11517-022-02523-1
90. Larpant N, Niamsi W, Noiphung J, Chanakiat W, Sakuldamrongpanich T, Kittichai V, et al. Simultaneous phenotyping of five Rh red blood cell antigens on a paper-based analytical device combined with deep learning for rapid and accurate interpretation. Anal Chim Acta. 2022;1207:339807. https://doi.org/10.1016/j.aca.2022.339807
91. Doan M, Sebastian JA, Caicedo JC, Siegert S, Roch A, Turner TR, et al. Objective assessment of stored blood quality by deep learning. Proc Natl Acad Sci. 2020;117(35):21381–90. https://doi.org/10.1073/pnas.2001227117
92. Kim E, Park S, Hwang S, Moon I, Javidi B. Deep learning-based phenotypic assessment of red cell storage lesions for safe transfusions. IEEE J Biomed Health Inform. 2022;26(3):1318–28. https://doi.org/10.1109/JBHI.2021.3104650
93. Quinn J, Campbell C, Gomez A, Kumar-Misir A, Watson S, Liwski D, et al. The successful implementation of an automated institution-wide assessment of hemoglobin and ABO typing to dynamically estimate red blood cell inventory requirements. Transfusion. 2019;59(7):2203–6. https://doi.org/10.1111/trf.15272
94. Guan L, Tian X, Gombar S, Zemek AJ, Krishnan G, Scott R, et al. Big data modeling to predict platelet usage and minimize wastage in a tertiary care system. Proc Natl Acad Sci U S A. 2017;114(43):11368–73. https://doi.org/10.1073/pnas.1714097114
95. Mirjalili M, Abouee-Mehrizi H, Barty R, Heddle NM, Sarhangian V. A data-driven approach to determine daily platelet order quantities at hospitals. Transfusion. 2022;62(10):2048–56. https://doi.org/10.1111/trf.17080
96. Li N, Arnold DM, Down DG, Barty R, Blake J, Chiang F, et al. From demand forecasting to inventory ordering decisions for red blood cells through integrating machine learning, statistical modeling, and inventory optimization. Transfusion. 2022;62(1):87–99. https://doi.org/10.1111/trf.16739
97. Schilling M, Rickmann L, Hutschenreuter G, Spreckelsen C. Reduction of platelet outdating and shortage by forecasting demand with statistical learning and deep neural networks: modeling study. JMIR Med Inform. 2022;10(2):e29978. https://doi.org/10.2196/29978
98. Sun X, Xu Z, Feng Y, Yang Q, Xie Y, Wang D, et al. RBC inventory-management system based on XGBoost model. Indian J Hematol Blood Transfus. 2021;37(1):126–33. https://doi.org/10.1007/s12288-020-01333-5
99. Fischer C, Bloder L, Neumüller C, Pimminger S, Affenzeller M, Winkler SM, et al. Prediction of blood demands in a hospital. In: *Proc EMSS 2010*. 2010. p. 7–12.
100. Rad J, Cheng C, Quinn JG, Abidi S, Liwski R, Abidi SSR. An AI-driven predictive modelling framework to analyze and visualize blood product transactional data for reducing blood products' discards. In: Michalowski M, Moskovitch R, editors. Artificial intelligence in medicine. Lecture notes in computer science. Cham: Springer International Publishing; 2020. p. 192–202. https://doi.org/10.1007/978-3-030-59137-3_18

101. Xiang RF, Quinn JG, Watson S, Kumar-Misir A, Cheng C. Application of unsupervised machine learning to identify areas of blood product wastage in transfusion medicine. Vox Sang. 2021;116:955–64. https://doi.org/10.1111/vox.13089
102. Huang X, Wang Y, Chen B, Huang Y, Wang X, Chen L, et al. Ability of a machine learning algorithm to predict the need for perioperative red blood cells transfusion in pelvic fracture patients: a multicenter cohort study in China. Front Med. 2021;8:694733. https://doi.org/10.3389/fmed.2021.694733
103. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337–40. https://doi.org/10.1038/s41591-019-0548-6
104. Sendak MP, D'Arcy J, Kashyap S, Gao M, Nichols M, Corey K, et al. A path for translation of machine learning products into healthcare delivery. EMJ Innov. 2020. https://doi.org/10.33590/emjinnov/19-00172
105. Mišić VV, Rajaram K, Gabel E. A simulation-based evaluation of machine learning models for clinical decision support: application and analysis using hospital readmission. NPJ Digit Med. 2021;4(1):1–11. https://doi.org/10.1038/s41746-021-00468-7
106. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. 2023;29(8):1930–40. https://doi.org/10.1038/s41591-023-02448-8
107. Center for Devices and Radiological Health. Good machine learning practice for medical device development: guiding principles. US Food Drug Administration. 2022 [cited 2023 Sept 13]. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles
108. Amrollahi F, Shashikumar SP, Holder AL, Nemati S. Leveraging clinical data across healthcare institutions for continual learning of predictive risk models. Sci Rep. 2022;12(1):8380. https://doi.org/10.1038/s41598-022-12497-7
109. Feng J, Phillips RV, Malenica I, Bishara A, Hubbard AE, Celi LA, et al. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. NPJ Digit Med. 2022;5(1):1–9. https://doi.org/10.1038/s41746-022-00611-y
110. Pressly MA, Parker RS, Waters JH, Beck SL, Jeyabalan A, Clermont G. Improvements and limitations in developing multivariate models of hemorrhage and transfusion risk for the obstetric population. Transfusion. 2021;61(2):423–34. https://doi.org/10.1111/trf.16216
111. McDermott MBA, Wang S, Marinsek N, Ranganath R, Foschini L, Ghassemi M. Reproducibility in machine learning for health research: still a ways to go. Sci Transl Med. 2021;13(586):eabb1655. https://doi.org/10.1126/scitranslmed.abb1655
112. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: an R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. Campbell Syst Rev. 2022;18(2):e1230. https://doi.org/10.1002/cl2.1230
113. Ho WH, Chang CS. Genetic-algorithm-based artificial neural network modeling for platelet transfusion requirements on acute myeloblastic leukemia patients. Expert Syst Appl. 2011;38(5):6319–23. https://doi.org/10.1016/j.eswa.2010.11.110
114. Aldaghi T, Morteza GH, Kargari M. Forecasting the amount of blood ordered in the obstetrics and gynaecology ward with the data mining approach. Indian J Hematol Blood Transfus. 2020;36(2):361–7. https://doi.org/10.1007/s12288-019-01203-9
115. Mitterecker A, Hofmann A, Trentino KM, Lloyd A, Leahy MF, Schwarzbauer K, et al. Machine learning–based prediction of transfusion. Transfusion. 2020;60(9):1977–86. https://doi.org/10.1111/trf.15935
116. Walczak S, Scharf JE. Transfusion cost containment for abdominal surgery with neural networks. Neural Process Lett. 2000;11(3):229–38. https://doi.org/10.1023/A:1009667711423
117. Walczak S, Scharf JE. Reducing surgical patient costs through use of an artificial neural network to predict transfusion requirements. Decis Support Syst. 2000;30(2):125–38. https://doi.org/10.1016/S0167-9236(00)00093-2
118. Walczak S, Mikhail E. Predicting estimated blood loss and transfusions in gynecologic surgery using artificial neural networks. Int J Healthc Inf Syst Inform. 2021;16:1–15. https://doi.org/10.4018/IJHISI.2021010101
119. Zhang J, Jiang L, Zhu X. A machine learning-modified novel nomogram to predict perioperative blood transfusion of Total gastrectomy for gastric cancer. Front Oncol. 2022;12:826760. https://doi.org/10.3389/fonc.2022.826760
120. Lammers D, Marenco C, Morte K, Conner J, Williams J, Bax T, et al. Machine learning for military trauma: novel massive transfusion predictive models in combat zones. J Surg Res. 2022;270:369–75. https://doi.org/10.1016/j.jss.2021.09.017
121. Kim KN, Kim DW, Jeong MA. The usefulness of a classification and regression tree algorithm for detecting perioperative transfusion-related pulmonary complications. Transfusion. 2015;55(11):2582–9. https://doi.org/10.1111/trf.13202
122. Murphree DH, Clifford L, Lin Y, Madde N, Ngufor C, Upadhyaya S, et al. Predicting adverse reactions to blood transfusion. In: *2015 International Conference on Healthcare Informatics*. IEEE; 2015. p. 82–9. https://doi.org/10.1109/ICHI.2015.17
123. Murphree DH, Clifford L, Lin Y, Madde N, Ngufor C, Upadhyaya S, et al. A clinical decision support system for preventing adverse reactions to blood transfusion. In: *2015 International Conference on Healthcare Informatics*. IEEE; 2015 p. 100–4. https://doi.org/10.1109/ICHI.2015.19
124. Murphree D, Ngufor C, Upadhyaya S, Madde N, Clifford L, Kor DJ, et al. Ensemble learning approaches to predicting complications of blood transfusion. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE; 2015. p. 7222–5. https://doi.org/10.1109/EMBC.2015.7320058
125. Ramoa A, Condeço J, Fdez-Riverola F, Lourenço A. HaemoKBS: a knowledge-based system for real-time, continuous categorisation of adverse reactions in blood recipients. Neurocomputing. 2021;423:756–67. https://doi.org/10.1016/j.neucom.2020.04.101
126. Borgmann DM, Mayr S, Polin H, Schaller S, Dorfer V, Obritzberger L, et al. Single molecule fluorescence microscopy and machine learning for rhesus D antigen classification. Sci Rep. 2016;6(1):32317. https://doi.org/10.1038/srep32317
127. Darlis D, Murwati H, Priramadhi RA, Ramdhani M, Nugraha MB. Functional simulation of human blood identification device using feed-forward artificial neural network for FPGA implementation. In: *2018 International Conference on Signals and Systems (ICSigSys)*. IEEE; 2018 p. 142–5. https://doi.org/10.1109/ICSIGSYS.2018.8372654
128. Ferraz A, Brito JH, Carvalho V, Machado J. Blood type classification using computer vision and machine learning. Neural Comput Applic. 2017;28(8):2029–40. https://doi.org/10.1007/s00521-015-2151-1

129. Rosales MA, de Luna RG. Computer-based blood type identification using image processing and machine learning algorithm. J Adv Comput Intell Intell Inform. 2022;26(5):698–705. https://doi.org/10.20965/jaciii.2022.p0698

How to cite this article: Maynard S, Farrington J, Alimam S, Evans H, Li K, Wong WK, et al. Machine learning in transfusion medicine: A scoping review. Transfusion. 2023. https://doi.org/10.1111/trf.17582

APPENDIX

MACHINE LEARNING METHOD CLASSIFICATION

In Table 2 of the main paper, we divide supervised machine learning methods into three broad groups: tree-based methods, neural networks, and other methods. Tree-based methods include classification and regression tree (CART), a single decision tree, and methods that train an ensemble of trees in parallel (e.g., random forests [RF]) or sequentially (e.g., gradient boosting machines, XGBoost, and CatBoost). Neural networks consist of layers of "neurons" inspired by biologic neurons and include fully connected neural networks (FCNNs), in which each "neuron" receives input from every "neuron" in the preceding layer, and alternative architectures that have been developed for different types of input data including convolutional neural networks (CNNs) for images and recurrent neural networks (e.g., the long short-term memory [LSTM] network) for sequences. Our final category includes any methods that are not based on decision trees and are not neural networks including generalized linear models (e.g., logistic regression and linear regression), support vector machines, naïve Bayes, K-nearest neighbors, and Markov chains.