

# Predicting Individual Patient Platelet Demand in a Large Tertiary Care Hospital Using Machine Learning

Merlin Engelke<sup>a, b</sup> Christian Martin Brieske<sup>c</sup> Vicky Parmar<sup>a, b</sup> Nils Flaschel<sup>a, b</sup>  
Anisa Kureishi<sup>a</sup> Rene Hosch<sup>a, b</sup> Sven Koitka<sup>a, b</sup> Cynthia Sabrina Schmidt<sup>c</sup>  
Peter A. Horn<sup>c</sup> Felix Nensa<sup>a, b</sup>

<sup>a</sup>University Medicine Essen, Institute for Artificial Intelligence in Medicine, Essen, Germany; <sup>b</sup>University Medicine Essen, Institute of Diagnostic and Interventional Radiology and Neuroradiology, Essen, Germany;  
<sup>c</sup>University Medicine Essen, Institute for Transfusion Medicine, Essen, Germany

## Keywords

Patient individual · Platelet prediction · Machine learning · Blood transfusion · Donor management · Platelet · Platelet concentrates

## Abstract

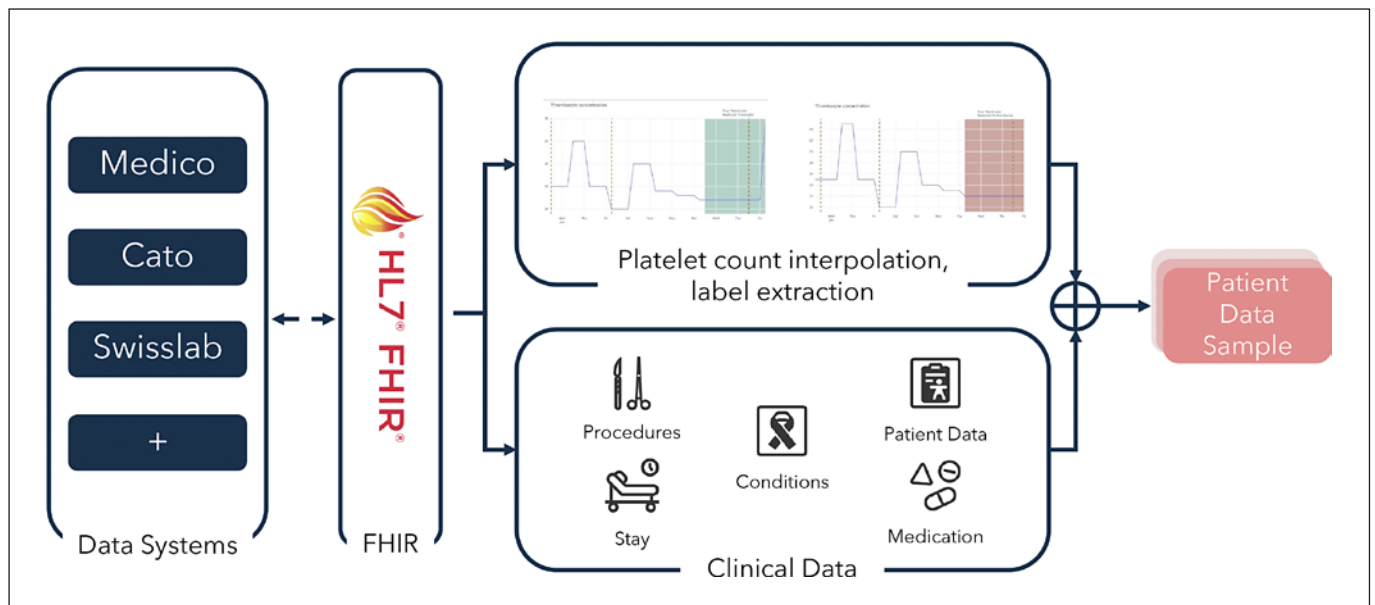
**Introduction:** An increasing shortage of donor blood is expected, considering the demographic change in Germany. Due to the short shelf life and varying daily fluctuations in consumption, the storage of platelet concentrates (PCs) becomes challenging. This emphasizes the need for reliable prediction of needed PCs for the blood bank inventories. Therefore, the objective of this study was to evaluate multimodal data from multiple source systems within a hospital to predict the number of platelet transfusions in 3 days on a per-patient level. **Methods:** Data were collected from 25,190 (42% female and 58% male) patients between 2017 and 2021. For each patient, the number of received PCs, platelet count blood tests, drugs causing thrombocytopenia, acute platelet diseases, procedures, age, gender, and the period of a patient's hospital stay were collected. Two models were trained on samples using a sliding window of 7 days as input and a day 3 target. The model predicts whether a patient will be transfused 3 days in the future. The model was trained with an excessive hyperparameter search using patient-level repeated 5-fold cross-validation to optimize the average macro F2-score. **Results:** The trained models were tested on 5,022 unique patients. The best-performing model has a

specificity of 0.99, a sensitivity of 0.37, an area under the precision-recall curve score of 0.45, an MCC score of 0.43, and an F1-score of 0.43. However, the model does not generalize well for cases when the need for a platelet transfusion is recognized. **Conclusion:** A patient AI-based platelet forecast could improve logistics management and reduce blood product waste. In this study, we build the first model to predict patient individual platelet demand. To the best of our knowledge, we are the first to introduce this approach. Our model predicts the need for platelet units for 3 days in the future. While sensitivity underperforms, specificity performs reliably. The model may be of clinical use as a pretest for potential patients needing a platelet transfusion within the next 3 days. As sensitivity needs to be improved, further studies should introduce deep learning and wider patient characterization to the methodological multimodal, multi-source data approach. Furthermore, a hospital-wide consumption of PCs could be derived from individual predictions.

© 2023 The Author(s).  
Published by S. Karger AG, Basel

## Introduction

It is an ever-increasing worldwide challenge to supply blood products to patients. In Germany, around 15,000 blood cell units are required every day, and about 45,000 blood products are transfused annually at our clinic, of which approximately 30,000 are red blood cell concen-



**Fig. 1.** Data source to patient data pipeline (taking multiple hospital systems into account) followed by the HL7 FHIR server, which extracts and stores the data, then the extraction process, and, finally, the per-patient storing of all relevant patients.

trates, 10,000 are platelet concentrates (PCs), and 5,000 are fresh frozen plasmas [1]. Furthermore, the ongoing demographic change is expected to increase the relative shortage of blood donors and thus increase the shortage of stored blood products [2–4]. Therefore, optimal use of this valuable resource is becoming increasingly important from a medical, ethical, and economic point of view. To solve these problems, improving logistics management in transfusion medicine is crucial.

Thus, coordinating the inventory, expiry, and consumption of PCs plays a vital role. A PC typically has a shelf life of 4 days after testing and screening procedures on the production day [5]. The transfusion of different patient collectives treated in a university hospital for hematology and oncology, transplantation, cardiac, and thoracic surgery leads to dramatic fluctuations in the consumption of PCs [6–8]. Therefore, storing PCs in line with requirements is challenging: on the one hand, a PC shortage is unacceptable, while on the other hand, it is necessary to minimize the wasteful disposal of blood donations.

According to the Paul Ehrlich Institute, 575,608 PCs were produced in Germany in 2020 [1]. From a total of 255,722 pool PCs, 48,467 PCs (19%) expired at the producer. From a total of 319,886 apheresis PCs, 22,411 (7%) expired at the producer [1]. Taking their market value into account, the disposal of unused PCs results in an economic loss of more than 30 million Euros per year. Thus, reducing the forfeiture rate by just two percent would save more than 5.5 million Euros annually [9].

The high expiration rates indicate that a patient-specific prediction of PC consumption is required to optimize logistics management. At present, however, the patient- and hospital-specific data already available in most hospitals in Germany are not sufficiently utilized for this purpose.

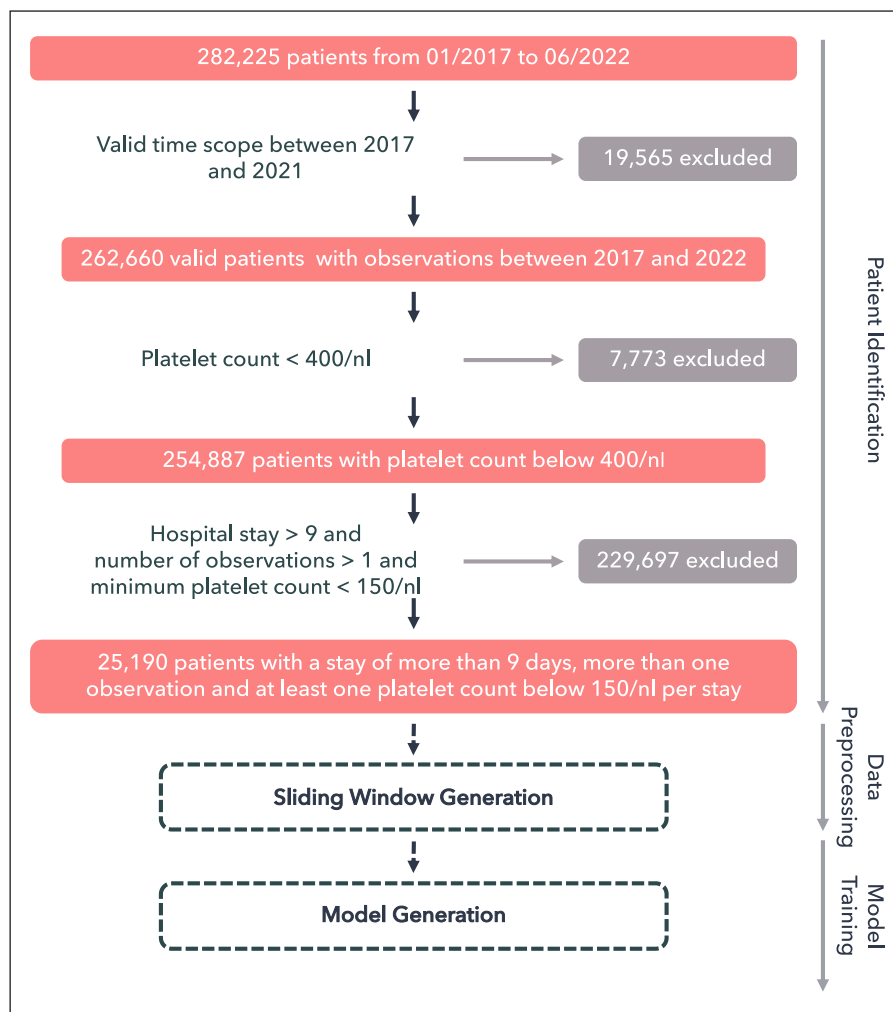
This retrospective study aimed to develop a novel AI-based approach to improve the prediction of individual PC demand. Because donation appointments must be scheduled and the donation products tested and screened, predictions for patient transfusion needs at the end of 3 days are critical. The objective of this study was, therefore, to create a machine learning-based model which predicts whether a patient requires a platelet transfusion 3 days into the future. This is achieved by extracting, transforming, and loading multimodal data from multiple source systems within the hospital.

## Materials and Methods

### Data Sources

The retrospective data were obtained from a hospital's internal Health Level Seven International (HL7) Fast Healthcare Interoperability Resources (FHIR) (<https://www.hl7.org/fhir>) server. The server is a data repository that stores information generated during the clinical routine at all hospital departments. It was queried for clinical and demographic information such as patient stay, laboratory testing, conditions, transfusions, medications, procedures, gender, and age. In addition, manual spot checks were performed to ensure data validity and quality between source systems and the extracted data set. At first, the completeness of extracted data be-

**Fig. 2.** Cohort identification process, which can be interpreted as a top-down flow diagram within the patient identification process. The entire cohort began with a total of 282,225 patients based on in-house platelet count observations between January 2017 and June 2022. The patient identification process can be divided into three filtering steps. The first filters the time scope of 2017–2022. The second filter excludes patients with platelet counts exclusively above 400 platelet/nL. The final filter ensures that a patient’s hospital stay is greater than 9 days, the number of blood observations is greater than one, and at least one platelet count observation needs to be below 150 platelet/nL. After the above-described filter, 25,190 patients are left for the scope of this study.



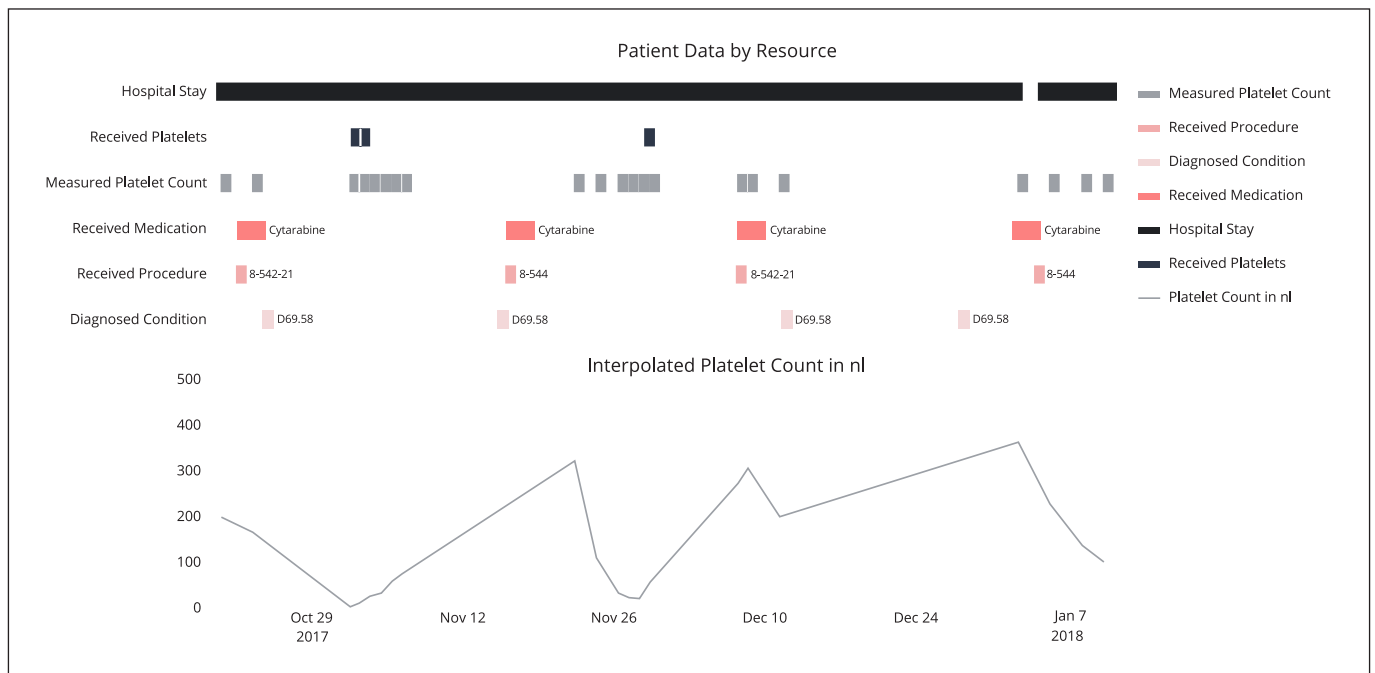
tween the FHIR server and training data was ensured by comparing the total number of extracted elements for each resource. Second, each FHIR resource’s data importers were checked for data completeness. Finally, to ensure end-to-end completeness, data mappings, as seen in Figure 3, were generated on a set of patients. Figure 1 exemplifies the data pipeline applied to extract the patient cohort. The following two subsections explain the data source in detail.

#### Subjects Identification and Preprocessing

The need for a PC donation is determined by the platelet count in the patient’s blood and specified in the hemotherapeutic guidelines [10, 11]. Our study included all patients, independent of the main or first diagnosis, who had platelet count observation between January 2017 and June 2022 and are in total 282,225 patients. A flow diagram visualizing the patient selection process in a schematic overview is represented in Figure 2.

First, a time scope filter was applied as only patients with observations between 2017 and 2021 were considered. While the normal laboratory range for PCs may vary slightly, the usual reference range is between 150 platelet/nL and 400 platelet/nL [12]. Therefore, patients with thrombocyte counts exclusively above the maximum reference range of 400 platelet/nL were excluded. Because each hospitalization had to pass another set of requirements to be included in the analysis, this stage of the identification process reduced the cohort size to 254,887 patients.

First, the hospital stay had to last at least 10 days. Patients with a shorter stay were considered out of scope as these patients had a very low upfront probability of needing a PC and typically had only very few data points. We found the probability for a patient to receive a PC during a stay between 0 and 9 days to be 0.34%. Second, patients with a single observation were considered out of scope because, according to transfusion guidelines, the platelet count should be measured before and after a patient receives a donation [11]. Finally, the minimum platelet count for one observation during a stationary stay had to be less than or equal to 150 platelet/nL, in accordance with the reference range for PC transfusions [12]. This filtering step yielded a final cohort size of 25,190 patients with valid hospitalizations to include in this study. For patients in the final cohort, we collected additional features such as procedures that may impact platelet function (e.g., extracorporeal circulation, hemodialysis), medications that frequently cause thrombocytopenia (e.g., chemotherapeutics such as cytarabine and gemcitabine), conditions that are known to cause thrombocytopenia either directly or through drug-based interference (e.g., leukemic and lymphoid malignancies), the number of received PC transfusions, and time scope of hospital stays. Figure 3 exemplifies the data map for one of the 25,190 patients, with all the occurring features from different resources. A detailed overview of the selected features from medications, procedures, and conditions can be found in Appendix A, available at [www.karger.com/doi/10.1159/000528428](http://www.karger.com/doi/10.1159/000528428).



**Fig. 3.** Exemplary data map of a cohort patient suffering from mantle cell lymphoma, which is represented on multiple occasions as a “diagnosed condition” with the ICD code “C83.1.” The patient was therefore subjected to a chemotherapeutic treatment plan with cytarabine (represented in four instances both as “received medication” and “received procedure” with the respective OPS codes “8-542-21” and “8-544”). This led to two occurrences of secondary thrombocytopenia (“D69.58”) with a minimal platelet count of 6/nL (November 1, 2017) and 24/nL (November 28, 2017) that were each treated with a PC transfusion (“received platelets”).

#### Sliding Window Sampling

The data needed to be transformed into equally dimensional samples to attain equally shaped data for machine learning models. Figure 4 shows a schematic overview of how each training sample’s sliding window was generated. The following criteria define a valid training sample: One platelet count observation needs to be present, and the minimum platelet count needs to be below 150 platelets/nL within the 7-day window.

Each input vector had a time scope of 7 days and a frequency of 12 h, thus 14 time slots. The window sampling was based on time logs of the patient encounters. Consequently, features for the respective time scope were collected using an aggregation function for each valid patient stay at the hospital.

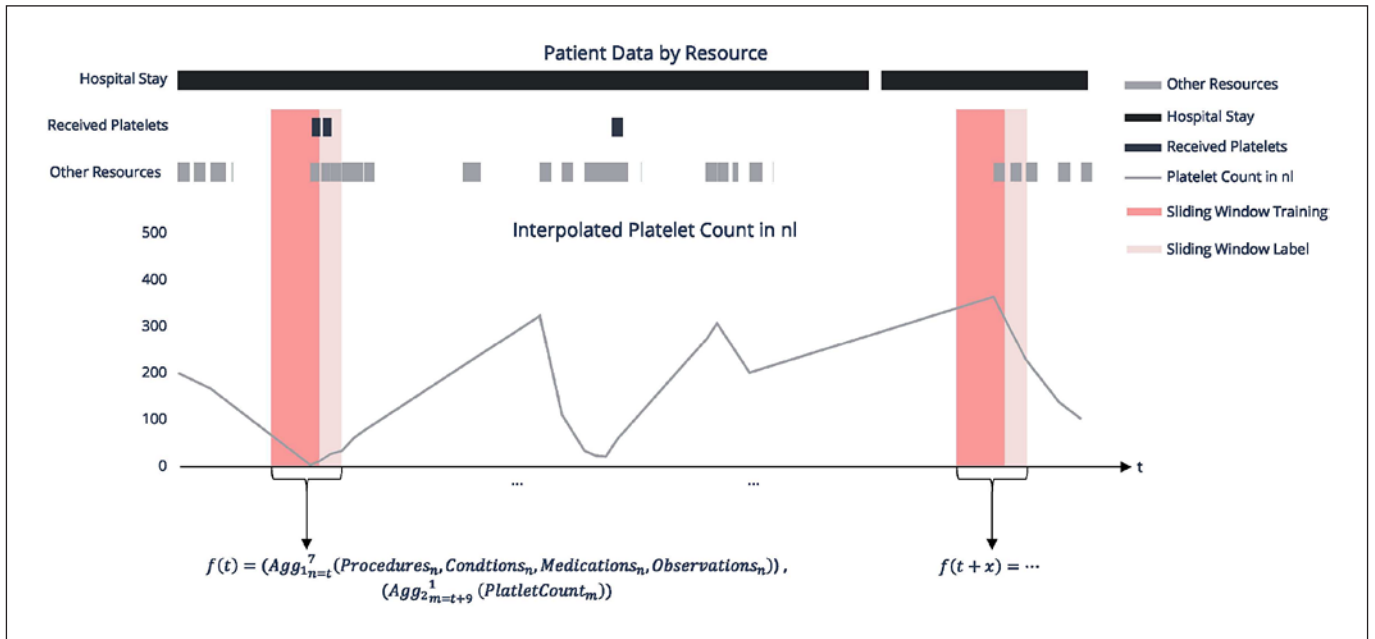
The aggregation function mapped each feature data event within each valid patient stay to one of the 14 time slots. Medications, conditions, and procedures were saved in a list per time slot. The most recent platelet count observation within a 12-h window was preserved, and subsequent empty spaces were linearly interpolated to the next platelet count observation. Additionally, it was recorded whether a platelet count value was interpolated if needed. Finally, each PC unit within a time slot was transformed into the sum of units within a day.

Time slot-independent metadata, such as age and gender, were stored for each patient. Finally, each target vector had a time scope of 1 day and a frequency of 24 h, which contained a list of consumed PC units 3 days into the future. Once preprocessing steps for one input and target vector were calculated, the sliding window algorithm stepped 1 day ahead, and the preprocessing re-started.

#### Platelet Binary Classification

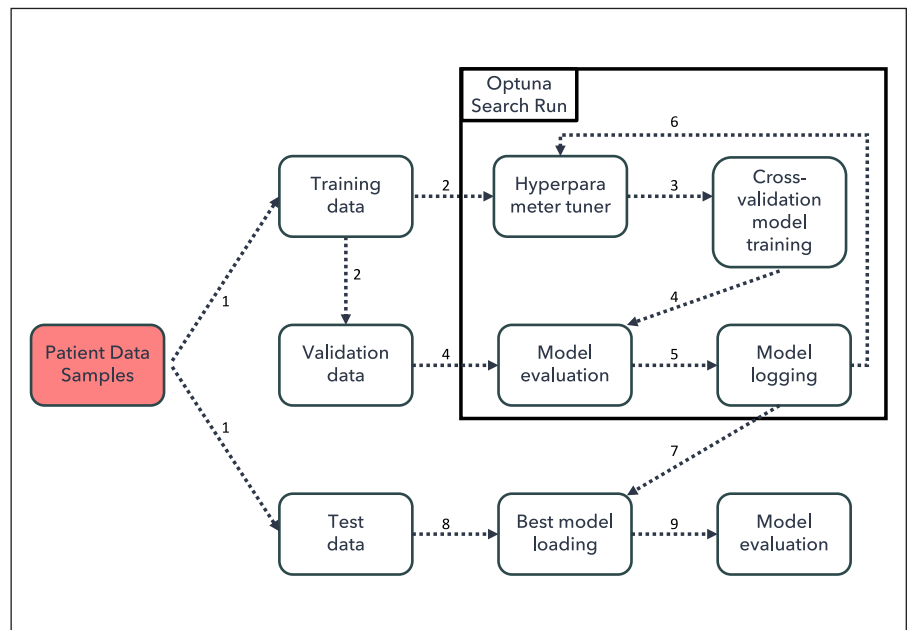
Classification algorithms were used as the decision of whether a patient should receive a PC can be translated into a binary problem. The first class (0) represented the case of no transfusion, and the second class (1) represented one or more transfusions. The data set was split on the patient level into training and test data using an 80/20 ratio. The data were highly zero-inflated as most samples did not receive any PC transfusion. To address the class imbalance (3.42%), all models were also trained using Gaussian noise up-sampling [13] and the kmeans\_SMOTE (synthetic minority over-sampling technique) [14] using default parameters. As the objective was to predict the third day, a single-output classifier was sufficient. For the model training and optimization, scikit-learn [15] [https://scikit-learn.org], eXtreme Gradient Boosting (XGBoost) [16] [https://xgboost.readthedocs.io], Random Forest [17], Dummy Classifier [18], and Optuna [19] [https://optuna.org] were used. Optuna was configured to run 1,000 trials for each model using a tree-structured Parzen estimator sampler [20] for hyperparameter sampling. Random sampling was used on the first 50 trials of each model as warm-up iterations for the tree-structured Parzen estimator algorithm. Samplers decrease the search time, and calculation cost of finding the best-performing hyperparameters for the chosen algorithm compared to a full grid search.

A median pruner (24) was used on top of the iterative process to stop the training process once the improvement plateaued. Pruning was disabled for the first 100 trials of each model run. All models were trained using fivefold stratified-group cross-validation, to ensure that cohorts within a split are not overlapping and the percentage of samples for each class is almost equal. An F2-score was used as an optimization metric to weigh sensitivity/recall



**Fig. 4.** Generation of the sliding window from raw data for training and labeling schematically. The sliding window frequency of  $x$  is one (day) with a frequency of 12 h and 7 days for each training sample and a frequency of 24 h and 3 days for each labeled sample.

**Fig. 5.** ML workflow using Optuna. First, the patient data samples were split into training and test data (1). The Optuna run was then initialized (2), which initialized the first algorithm and prepared the data for a cross-validated run (3). In cross-validation, the data were further split into the desired amount of data sets (which consist of training and validation data sets) to further prevent overfitting and selection bias. Each set was then trained, and overall folds were averaged to calculate the model's predictive performance on the validation set (4). The model was then logged with its performance scores and hyperparameter (5). If the pruner decided to train another model, the process would start again at the hyperparameter tuning within the given search space (6). Once the evaluation was finished, the parameters of the best-performing model could be loaded and evaluated on the test data (7–9).



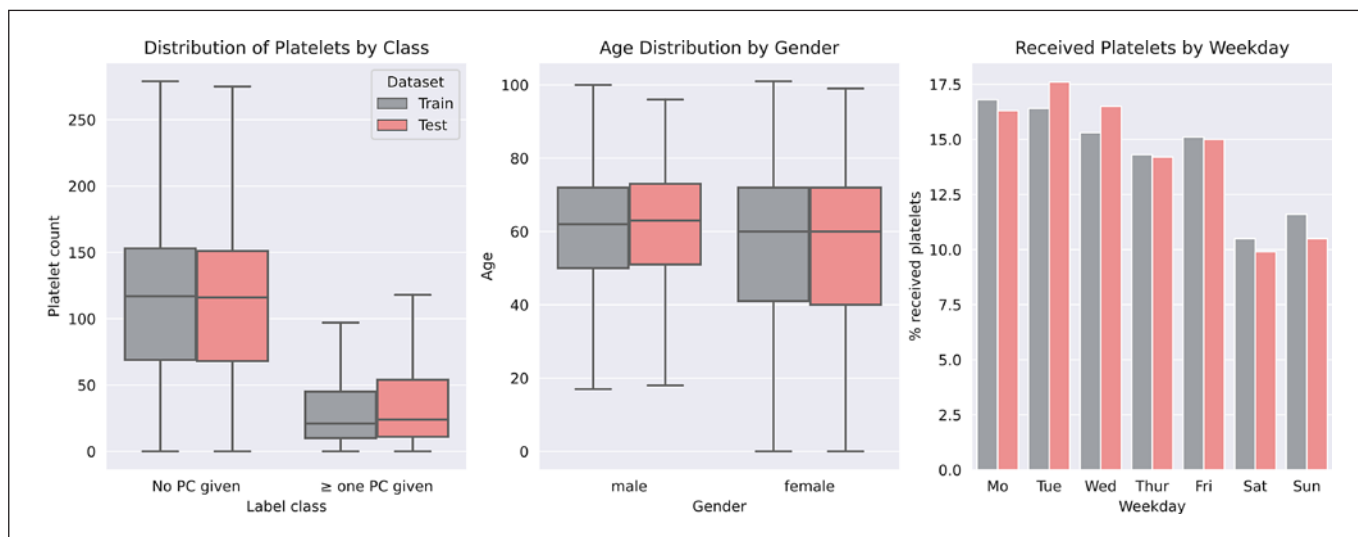
higher than precision. Figure 5 depicts the training process for the machine learning models in a schematic overview.

A dummy classifier with a stratified strategy was used as a baseline model to compare against the above-described classifiers. Dummy classifiers disregard the input features and solely depend on the output values. The stratified method considered the a priori probability and was, therefore, a valid baseline classifier. Subsequently, an XGBoost and a Random Forest model were trained excessively with the hyperparameter search spaces seen in Table 1 using the ML workflow described in Figure 5.

## Results

### Patient Characteristics

After the processing steps seen in Figures 2 and 4, the cohort consisted of 25,190 patients with a hospital stay of at least 10 days between 2017 and 2021. The cohort data were randomly split on the patient level into 80/20 subsets for training and test data sets. In the training set, the patient's ages ranged from 0 to 101 years, and the mean pa-



**Fig. 6.** Patient characteristics by training and test sets. The figure on the left-hand side shows the distribution of platelets by data class. The left-hand figure indicates that patients in the first label class have, in most cases, a platelet count below 50/nL. The center figure shows the age distribution by gender. At the lower end of the distribution, the boxplot has an outlier cutoff for the male cohort below 19 years. However, this does not mean that there are no male patients younger than 19 years. The figure on the right-hand side illustrates the relative received platelets by weekday. It can be observed that most patients receive transfusions between Monday and Friday.

**Table 1.** Hyperparameter spaces for XGBoost and Random Forest

Parameter	Search space
<b>XGBoost with gbtrees booster</b>	
Objective	Binary: logistic
Eval_metric	AUCPR <sup>a</sup>
Lambda	10 <sup>-8</sup> -1.0
Alpha	10 <sup>-8</sup> -1.0
Max_depth	10-30
Eta	10 <sup>-8</sup> -1.0
Gamma	10 <sup>-8</sup> -1.0
Grow_policy	Depthwise, lossguide
<b>XGBoost with gbtrees dart</b>	
Objective	Binary: logistic
Eval_metric	AUCPR <sup>a</sup>
Lambda	10 <sup>-8</sup> -1.0
Alpha	10 <sup>-8</sup> -1.0
Max_depth	10-30
Eta	10 <sup>-8</sup> -1.0
Gamma	10 <sup>-8</sup> -1.0
Grow_policy	Depthwise, lossguide
Sample_type	Uniform, weighted
Normalize_type	Tree, forest
Rate_drop	10 <sup>-8</sup> -1.0
Skip_drop	10 <sup>-8</sup> -1.0
Criterion	Gini, entropy
<b>Random Forest</b>	
N_estimators	100-1,000
Splitter	Best, random
Max_depth	10-30
Max_depth	10-30

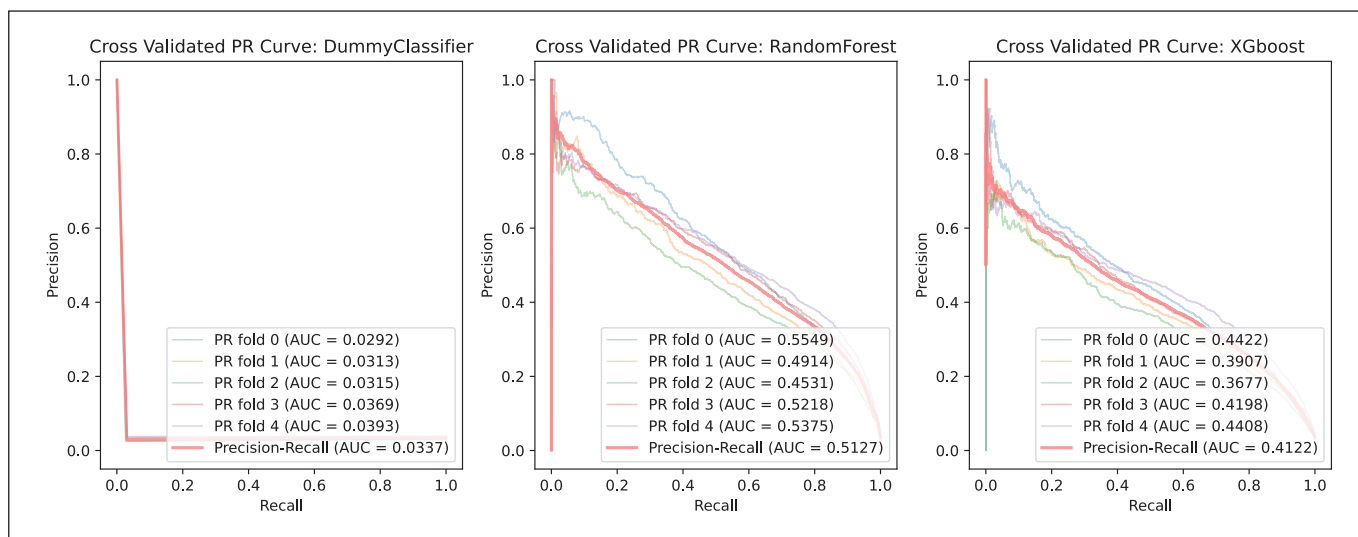
<sup>a</sup> AUCPR, area under the precision-recall curve.

tient age was  $56.5 \pm 21.8$  years. 42% of the patients were female, and 58% were male. In the test set, the patient age ranged from 0 to 99 years, and the mean patient age was  $56.9 \pm 22$  years. The gender distribution was the same as in the training set. Figure 6 illustrates the data set distribution from different perspectives.

#### Model Performance

During the observed period of 5 years, 25,190 patients were transfused, with a total of 54,473 PC units. Optuna was configured to run for 1,000 trials for the XGBoost and Random Forest models. The performance of the best-performing classification models on the training data is shown in Table 2. All metrics (area under the precision-recall curve [AUCPR] score, F1 score, MCC score, precision, specificity, and sensitivity) could be improved compared to the dummy classifier. However, all models lack sensitivity.

Figure 7 shows the AUCPR curve across all trained models. The Dummy Classifier is displayed as an almost horizontal line with a score of 0.03, suggesting that the model has no discriminative ability to diagnose patients with and without needing a PC transfusion. The slope of the AUCPR curve on the Random Forest and XGBoost models indicates that the model performs better than the Dummy Classifier.



**Fig. 7.** Area under the precision-recall curve (AUCPR) with cross-validation for each of the five folds. The steepness and roundness toward the right upper corner represent good performance as it is the ideal maximum true-positive-to-false-negative ratio. Furthermore, it clearly shows that the AUCPR scores are almost equal as the data are evenly distributed across every data set created from K-fold cross-validation.

**Table 2.** Models performance on cross-validation and test data for the Dummy Classifier, XGBoost, and Random Forest

	Model					
	AUCPR	MCC	F1-score	Precision	Specificity	Sensitivity
<i>Cross-validation</i>						
Dummy Classifier	0.0339	0.0006	0.0352	0.0354	0.9671	0.0351
XGBoost	0.4100	0.3939	0.4044	0.4906	0.9875	0.3441
XGBoost with GNUS	0.2618	0.3507	0.3688	0.4083	0.9830	0.3364
XGBoost with SMOTE	0.4435	0.4036	0.4067	0.5348	0.99	0.3284
Random Forest	0.5131	0.3968	0.3596	0.6809	0.9960	0.2445
Random Forest with GNUS	0.51	0.3947	0.3558	0.6843	0.9961	0.241
Random Forest with SMOTE	0.5166	0.3966	0.3565	0.6908	0.9963	0.2407
<i>Test set</i>						
Dummy Classifier	0.0344	0.0018	0.0422	0.0433	0.9662	0.0413
XGBoost	0.4497	0.4247	0.4335	0.5342	0.9882	0.3648
XGBoost with GNUS	0.2916	0.3762	0.3939	0.4407	0.9833	0.3563
XGBoost with SMOTE	0.4786	0.4252	0.4261	0.5715	0.9905	0.3397
Random Forest	0.5467	0.4281	0.3890	0.7267	0.9963	0.2656
Random Forest with GNUS	0.5499	0.4223	0.3824	0.7228	0.9963	0.2600
Random Forest with SMOTE	0.5497	0.4232	0.3824	0.7270	0.9964	0.2594

GNUS, Gaussian noise up-sampling; SMOTE, synthetic minority over-sampling technique.

## Discussion/Conclusion

In the present study, we evaluated multimodal data from multiple source systems in a hospital to form a prediction of PC transfusions in 3 days on a per-patient level. To the best of our knowledge, this is the first study to extract individual patient data to decide whether a patient needs a PC transfusion in 3 days for a cohort of more than 25,000 patients on a per-patient level.

XGBoost was the best-performing model with the highest score. However, the recall was too low to integrate the model into a live system as many true positives were misclassified. Yet, specific detection of true negatives was achieved. Therefore, clinical implementation of the trained models could be considered a pretest. The model could run in the background of the clinical routine and filter out all patients who do not need a PC transfusion.

Other studies have investigated the prediction of the blood product requirements for clinic-wide consumption. These studies have already examined the impact of machine learning models on hospital-wide platelet demand [21–24] and have found that it can be predicted with high accuracy [21, 23, 24]. Schilling et al. [21] found that their models could reduce both the shortage and waste of PCs within their institution at RWTH Aachen University Hospital. One limitation of previous studies in this field was the insufficient sensitivity to outliers. Outliers are days of extremely high or low demand for PCs. A patient-specific model, on the other hand, could identify days with outliers of PC demand and react prospectively toward days with a PC shortage.

The current high expiration rates of platelets are the driving factor behind the potential use in clinical routine. Usually, PCs are bought from a third party, supplied by an in-house institution, or both. However, PC demand is determined by the transfusing physicians and often – as is the case at our institution – not reported back to the provider. Digitalization of hospital infrastructure is a potential resource investment in patient care. Concerning PC demand, a system for predicting patient PC transfusion needs could lead to reduced PC expiry, and a doctor-independent decision process could consequently result in widespread economic benefits.

Furthermore, a model that predicts demand at the individual patient level would be an essential step toward improving patient blood management. Patient blood management is a central hemotherapeutic concept for improving patient safety that aims at reducing the need for blood transfusions and takes hospital-wide blood product management into account [25]. By improving the prediction of individual PC supply through our model, potential PC shortages could be prevented.

The most significant limitation of our model is its low sensitivity, which might partly be the result of missing data, particularly regarding future events such as crucial planned medical interventions (e.g., stem cell transplantation). Furthermore, data on medical appliances booked for surgery, especially heart and lung support devices, are currently not stored digitally in our institution. Thus, a new framework for integrating these data into our system will have to be constructed. In addition, the machine learning models used in our study could be considered relatively trivial from a technical perspective. However, a deep learning-based model may improve the model's performance and should be tested in further studies. A recurrent neural network or transformer network with a multi-input, multi-output architecture will be designed. Furthermore, diagnoses and procedures are not permanently recorded for each patient in a timely manner and also only for accounting purposes, so the integration in a live scenario of the data extraction process has to be adapted

to using an identifier that merely extracts data of future events in FHIR.

In this study, we built the first model to predict patient individual PC demand. We have shown that an individual patient prediction of patients' PC transfusion needs can be ascertained using various data sources within the FHIR ecosystem and by applying traditional classification algorithms. However, further studies should apply this methodological data approach and adapt the machine learning models to more sophisticated methods, like deep neural networks, to improve sensitivity scores. In particular, however, the validity of a machine learning model depends on the scope and quality of the available data used as input.

In the future, great improvements in the predictive validity of corresponding models can be expected through a consistent, timely, and structured collection of all relevant healthcare data in the clinical routine. Once implemented, predictions for clinic-wide consumption may be derived, laying the groundwork for more efficient, ethical, and economical PC management.

### Statement of Ethics

This study was conducted in compliance with the Ethics Committee of the Medical Faculty of the University Duisburg-Essen, approval number 20-9386-BO. Due to the retrospective nature of the study, the requirement of written informed consent was waived by the Ethics Committee.

### Conflict of Interest Statement

The authors have no conflicts of interest to declare.

### Funding Sources

The funding was based on a resolution of the German Bundestag by the Federal Government (Az. ZMVI1-2519DAT713).

### Author Contributions

Merlin Engelke and Vicky Parmar acquired the data. Felix Nensa, Peter Alexander Horn, Merlin Engelke, Vicky Parmar, Rene Hosch, and Sven Koitka designed the study, analyzed the data, and co-wrote the manuscript. Cynthia Sabrina Schmidt, Nils Flaschel, Christian Martin Brieske, and Anisa Kureishi contributed to the study design and critically revised the manuscript. All the authors approved its final content.

### Data Availability Statement

The data supporting this study's findings are available from the corresponding author upon reasonable request.



## References

- 1 Henseler O. Bericht zur Meldung nach § 21 TFG für das Jahr 2020; 2021. [cited 2022 Jul 8]. Available from: [https://www.pei.de/SharedDocs/Downloads/DE/regulation/meldung/21-tfg/21-tfg-berichte/2020-tfg-21-bericht.pdf?\\_\\_blob=publicationFile&v=4](https://www.pei.de/SharedDocs/Downloads/DE/regulation/meldung/21-tfg/21-tfg-berichte/2020-tfg-21-bericht.pdf?__blob=publicationFile&v=4).
- 2 Greinacher A, Fendrich K, Hoffmann W. Demographic changes: the impact for safe blood supply. *Transfus Med Hemother*. 2010;37(3):141–8.
- 3 Greinacher A, Weitmann K, Schönborn L, Alpen U, Gloger D, Stangenberg W, et al. A population-based longitudinal study on the implication of demographic changes on blood donation and transfusion demand. *Blood Adv*. 2017 Jun;1(14):867–74.
- 4 University of Economics in Katowice, Faculty of Management, Twaróg S, Sołtysek J, University of Economics in Katowice, Faculty of Management, Majewska J, University of Economics in Katowice, Faculty of Informatics and Communication, et al. Influence of demographic change on the blood services in Poland: logistics as a remedy for the future. *Gospod Mater Logistyka*. 2019 Mar;2019(3):2–11.
- 5 Festlegung der Haltbarkeitsfrist von Thrombozytenkonzentraten mit dem Ziel der Reduktion lebensbedrohlicher septischer Transfusionsreaktionen durch bakterielle Kontamination. *Votum Des Arbeitskreises Blut*. 2008 Dec;51(12):1484.
- 6 Hamada SR, Garrigue D, Nogue H, Meyer A, Boutonnet M, Meaudre E, et al. Impact of platelet transfusion on outcomes in trauma patients. *Crit Care*. 2022 Dec;26(1):49.
- 7 Yanagawa B, Ribeiro R, Lee J, Mazer CD, Cheng D, Martin J, et al. Platelet transfusion in cardiac surgery: a systematic review and meta-analysis. *Ann Thorac Surg*. 2021 Feb;111(2):607–14.
- 8 Wandt H, Schäfer-Eckart K, Greinacher A. Platelet transfusion in hematology, oncology and surgery. *Dtsch Arztebl Int*. 2014 Nov;111(48):809–15.
- 9 DRK-BSD West. *Preisliste des DRK-BSD West vom 1.1.*; 2019.
- 10 Kaufman RM, Djulbegovic B, Gernsheimer T, Kleinman S, Tinmouth AT, Capocelli KE, et al. Platelet transfusion: a clinical practice guideline from the AABB. *Ann Intern Med*. 2015 Feb;162(3):205–13.
- 11 *Querschnitts-Leitlinien Hämotherapie*; 2020: 287.
- 12 Daly ME. Determinants of platelet count in humans. *Haematologica*. 2011 Jan;96(1):10–3.
- 13 Branco P, Torgo L, Ribeiro RP. A survey of predictive modeling on imbalanced domains. *ACM Comput Surv*. 2016 Nov;49(2):1–50.
- 14 Douzas G, Bacao F, Last F. Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. *Inf Sci*. 2018 Oct;465:1–20.
- 15 Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12:2825–30.
- 16 Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. Xgboost: extreme gradient boosting. *R Package Version*. 2015;1(4):1–4.
- 17 sklearn.ensemble.RandomForestClassifier [Internet]. Scikit-Learn. [cited 2022 Jul 19]. Available from: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>.
- 18 sklearn.dummy.DummyClassifier [Internet]. Scikit-Learn. [cited 2022 Jul 19]. Available from: <https://scikit-learn.org/stable/modules/generated/sklearn.dummy.DummyClassifier.html>.
- 19 Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: a next-generation hyperparameter optimization framework proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining; 2019:2623–31.
- 20 optuna.pruners.MedianPruner — Optuna 2.10.1 documentation [Internet]. [cited 2022 Jul 19]. Available from: <https://optuna.readthedocs.io/en/stable/reference/generated/optuna.pruners.MedianPruner.html>.
- 21 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 Feb;10(2):e29978.
- 22 Perelman I, Fergusson D, Lampron J, Mack J, Rubens F, Giulivi A, et al. Exploring peaks in hospital blood component utilization: a 10-year retrospective study at a large multisite academic centre. *Transfus Med Rev*. 2021 Jan;35(1):37–45.
- 23 Motamedi M, Li N, Down DG, Heddle NM. Demand forecasting for platelet usage: from univariate time series to multivariate models; 2021 Jan. [cited 2022 Mar 7]. Available from: <https://arxiv.org/abs/2101.02305v1>.
- 24 Fanoody B, Malmir B, Jahantigh FF. Reducing demand uncertainty in the platelet supply chain through artificial neural networks and ARIMA models. *Comput Biol Med*. 2019 Oct;113:103415.
- 25 Leahy MF, Hofmann A, Towler S, Trentino KM, Burrows SA, Swain SG, et al. Improved outcomes and reduced costs associated with a health-system-wide patient blood management program: a retrospective observational study in four major adult tertiary-care hospitals. *Transfusion*. 2017;57(6):1347–58.