Bleeding Risk Assessment in End-Stage Kidney Disease: Validation of Existing Risk Scores and Evaluation of a Machine Learning-Based Approach

Stephan Nopp[1] Clemens P. Spielvogel[2,3] Sabine Schmaldienst[4] Renate Klauser-Braun[5] Matthias Lorenz[6] Benedikt N. Bauer[1] Ingrid Pabinger[1] Marcus Säemann[7] Oliver Königsbrügge[1] Cihan Ay[1]

[1] Clinical Division of Haematology and Hemostaseology, Department of Medicine I, Medical University of Vienna, Vienna, Austria
[2] Division of Nuclear Medicine, Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Vienna, Austria
[3] Christian Doppler Laboratory for Applied Metabolomics, Vienna, Austria
[4] Department of Medicine I, Clinic Favoriten, Vienna, Austria
[5] Department of Medicine III, Clinic Donaustadt, Vienna, Austria
[6] Vienna Dialysis Centre, Vienna, Austria
[7] Department of Medicine VI, Clinic Ottakring, Vienna, Austria

Address for correspondence Cihan Ay, MD, Clinical Division of Haematology and Haemostaseology, Department of Medicine I, Medical University of Vienna, Waehringer Guertel 18-20, A-1090 Vienna, Austria (e-mail: cihan.ay@meduniwien.ac.at).

Thromb Haemost 2022;122:1558–1566.

Abstract

Background Patients with end-stage kidney disease (ESKD) on hemodialysis (HD) are at increased risk for bleeding. However, despite relevant clinical implications regarding dialysis modalities or anticoagulation, no bleeding risk assessment strategy has been established in this challenging population.

Methods Analyses on bleeding risk assessment models were performed in the population-based Vienna InVestigation of Atrial fibrillation and thromboemboLism in patients on hemoDialysIs (VIVALDI) study including 625 patients. In this cohort study, patients were prospectively followed for a median observation period of 3.5 years for the occurrence of major bleeding. First, performances of existing bleeding risk scores (i.e., HAS-BLED, HEMORR$_2$HAGES, ATRIA, and four others) were evaluated in terms of discrimination and calibration. Second, four machine learning-based prediction models that included clinical, dialysis-specific, and laboratory parameters were developed and tested using Monte Carlo cross-validation.

Results Of 625 patients (median age: 66 years, 37% women), 89 (14.2%) developed major bleeding, with a 1-year, 2-year, and 3-year cumulative incidence of 6.1% (95% confidence interval [CI]: 4.2–8.0), 10.3% (95% CI: 8.0–12.8), and 13.5% (95% CI: 10.8–16.2), respectively. C-statistics of the seven contemporary bleeding risk scores ranged between 0.54 and 0.59 indicating poor discriminatory performance. The HAS-BLED score showed the highest C-statistic of 0.59 (95% CI: 0.53–0.66). Similarly, all four machine learning-based predictions models performed poorly in internal validation (C-statistics ranging from 0.49 to 0.55).

Conclusion Existing bleeding risk scores and a machine learning approach including common clinical parameters fail to assist in bleeding risk prediction of patients on HD.

Keywords
- renal anemia
- hemodialysis
- risk assessment
- hemorrhage
- atrial fibrillation

received
November 8, 2021
accepted after revision
January 27, 2022
published online
June 19, 2022

DOI https://doi.org/10.1055/a-1754-7551.
ISSN 0340-6245.

Therefore, new approaches, including novel biomarkers, to improve bleeding risk prediction in patients on HD are needed.

Introduction

Chronic kidney disease, and especially end-stage kidney disease (ESKD), is associated with an increased risk for bleeding. The bleeding tendency in patients with ESKD is multifactorial and influenced by dysfunctional platelets, impaired platelet–vessel wall interactions, renal anemia, and treatment effects including anticoagulation therapy.[1] The bleeding rate in patients on hemodialysis (HD) has been reported to be 9 per 100 patient-years and 15% of the bleeds result in death.[2–8] Stratification of bleeding risk could be useful to individualize patient management; however, no valid approach to assess the risk of bleeding in this challenging patient population has been established and recommended to date.

Existing bleeding risk assessment tools, such as the HAS-BLED (Hypertension, Abnormal renal/liver function, Stroke, Bleeding history or predisposition, Labile international normalized ratio, Elderly, Drugs/alcohol concomitantly) score, are widely used in nondialysis patients and have been incorporated in several guidelines to assess bleeding risk in patients on antithrombotic therapy, particularly in patients with atrial fibrillation (AF) on oral anticoagulation.[9–12] Optimal bleeding risk assessment in patients on dialysis is desirable for several reasons during patient management: (1) bleeding risk could be considered to enable informed decision making such as the ideal dialysis treatment modality as several studies highlighted significantly lower bleeding rates in patients on peritoneal dialysis compared with HD.[13–15] (2) High-risk patients might benefit from close monitoring, control of risk factors, and decision of anticoagulation regimens for HD. (3) Further, AF is common in patients on HD and oral anticoagulation for stroke prevention is indicated in the vast majority of ESKD patients with AF. However, studies on the benefit of anticoagulation in dialysis patients are inconclusive and excess risk of bleeding during vitamin K antagonist treatment is observed.[16–19] Therefore, bleeding risk assessment has the potential to identify patients eligible for anticoagulation in the AF population and guide patient management in the general dialysis population.

Thus, we aimed at improving bleeding risk prediction in HD patients using a two-step approach. First, we evaluated known bleeding risk assessment tools in a prospective and observational study of patients with ESKD on HD; second to the standard statistical models, we developed four machine learning-based prediction models to identify novel approaches and predictors for bleeding risk assessment.

Methods

Study Design, Population, and Outcome

The Vienna InVestigation of AtriaL Fibrillation and thromboembolism in hemoDIalysis (VIVALDI) study is a prospective multicenter population-based cohort study, in which 625 patients were recruited in a cross-sectional fashion between April 2014 and July 2015 at seven dialysis centers in Vienna, Austria. All adult patients requiring chronic HD except patients who were pregnant or lactating were eligible for inclusion. A detailed description of the study design and procedures has been published elsewhere.[16,20] The VIVALDI study was approved by the local ethics committees (No. 1146/2014) and all patients provided written informed consent.

Patient demographics (age, gender, body mass index, smoking status) and clinical characteristics (primary kidney disease, AF, diabetes, history of bleeding or stroke, and other comorbidities) were assessed in a structured interview verified against medical records. Dialysis-specific parameters (e.g., remaining diuresis, ultrafiltration rate) and laboratory measurements were extracted from the electronic medical documentation system. After inclusion, patients were prospectively followed for a maximum of 1,350 days for the occurrence of major bleeding as defined by the International Society on Thrombosis and Hemostasis (ISTH)[21] and all-cause death. All outcomes were adjudicated by independent experts upon chart review and based on imaging evidence or autopsy findings.

In this prediction model validation and development study, we adhered to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement.[22]

Bleeding Risk Assessment Tools

Currently existing bleeding risk scores were selected from recent reviews on bleedings risk assessment tools in various diseases.[23–25] Seven bleeding risk prediction scores potentially applicable to the HD setting were identified and evaluated. The HAS-BLED,[26] ATRIA,[27] ORBIT,[28] OBRI,[29] and mORBI[30] scores were calculated based on the variables collected at study inclusion for each patient as defined in their original development studies. For the HEMORR$_2$HAGES risk score, two variables (i.e., "excessive fall risk" and "genetic factors") could not be considered due to the lack of baseline information with respect to fall risk and CYP2C9 single nucleotide polymorphisms. Further, history of bleeding was used as a proxy for "recent bleeding" for the score developed by Shireman et al.[31] A detailed explanation on the calculation of each risk score is available in the ►**Supplementary Methods** (available in the online version).

Statistical Analysis

Patient characteristics are presented as frequencies (percentage), mean ($\pm$ standard deviation), or median (interquartile range), as appropriate. Follow-up time was calculated using the reverse Kaplan–Meier method and bleeding rates are presented using a Poisson model as well as cumulative

incidences at 1, 2, and 3 years after study inclusion. We accounted for death and renal transplantation as a competing risk.

Performance of the bleedings risk scores was assessed using discrimination and calibration. Discriminatory ability of each risk score was evaluated using C-statistics including 95% confidence intervals (CIs). Calibration is presented visually using calibration plots comparing predicted versus actual bleeding risks. In addition, the Brier score, which is a measure for accuracy of probabilistic predictions, was calculated for each bleeding risk model. Further, net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were computed for each risk score when compared with the HAS-BLED score. Missing values (displayed in ►**Table 1**) were imputed using single imputation by predictive mean matching. Statistical analysis was conducted using R (version 3.6.2; R Core Team, 2019) and the following add-on packages were used: survival, cmprsk, riskRegression, pec, prodlim, epitools, predictABEL, and mice.

Machine Learning Methods

Four machine learning algorithms were employed for the binary classification of whether a patient experienced major bleeding within 24 months after study inclusion. Patients who were lost to follow-up or received renal transplantation were excluded from the subsequent analysis. The development of the supervised models consisted of three steps: data preprocessing, model training, and model validation (►**Fig. 1**). In the data-preprocessing step, handling of missing patient characteristics and laboratory parameters was conducted using a k-nearest neighbor imputation approach. Further, potential issues arising from class distribution imbalances were mitigated using a combination of random undersampling and the synthetic minority oversampling technique (SMOTE).[32] Both, imputation and resampling were performed after splitting into training and test set. Resampling was only performed on the training data. The k-nearest neighbor model for imputation was trained based on the training data and applied to training and test set to avoid data leakage. The training was performed

Table 1 Patient characteristics

	Total study cohort ($n = 625$)	Subgroup of patients with atrial fibrillation ($n = 165$)
Demographics		
Age, y	66 (54.5–75.0)	72 (64.0–78.0)
Male sex	394 (63.0%)	116 (70.3%)
BMI, kg/m^2 (*2*)	25.7 (22.4–29.4)	25.8 (22.8–29.6)
Etiology of end-stage kidney disease		
Diabetic nephropathy	160 (25.6)	43 (26.1)
Vascular nephropathy	121 (19.4)	39 (23.6)
Glomerulonephritis	81 (13.0)	16 (9.7)
Atrophic nephropathy	57 (9.1)	16 (9.7)
Cystic nonhereditary nephropathy	36 (5.8)	10 (6.1)
Hereditary nephropathy	31 (5.0)	5 (3.0)
Nephrectomy	20 (3.2)	10 (6.1)
Toxic nephropathy	28 (4.5)	9 (5.4)
Other causes	91 (14.6)	17 (10.3)
Dialysis-specific parameters		
Cumulative time on hemodialysis, y	2.7 (1.0–5. 0)	3.0 (1.1–6.0)
Ultrafiltration rate, mL (35)	2000 (982.5–3,000)	2100 (997.5–3,000)
Remaining diuresis, mL/d (24)	500 (0–1,000)	325 (0–1,000)
History of kidney transplantation	90 (14.4)	26 (15.8)
History of peritoneal dialysis (15)	46 (7.5)	11 (6.7)
Comorbidities		
Atrial fibrillation	165 (26.4%)	165 (100%)
Diabetes mellitus (2)	237 (38.0)	69 (42.1)
Heart failure	183 (29.3)	70 (42.4)
Coronary artery disease	232 (37.1)	83 (50.3)
Peripheral artery disease	197 (31.5)	57 (34.6)
Artificial heart valve	43 (6.9)	21 (12.7)
Hypertension	574 (91.8)	152 (92.1)

Table 1 (*Continued*)

Active or history of cancer[a]	152 (24.3)	56 (33.9)
Prior stroke/TIA/systemic embolism	137 (21.9)	48 (29.1)
Prior clinically relevant bleeding	147 (23.5)	51 (30.9)
Prior major bleeding	67 (10.7)	21 (12.7)
Prior myocardial infarction	104 (16.6)	35 (21.1)
Prior venous thromboembolism	61 (9.8)	25 (15.2)
Current or past smoker (*11*)	305 (49.7)	75 (46.3)
CHA_2DS_2-VASc score	3 (2–5)	4 (3–5)
HAS-BLED score	3 (2–4)	4 (3–5)
Laboratory parameters		
Hemoglobin, g/dL (*27*)	10.9 (10.0–11.7)	11.1 (10.1–11.9)
Hematocrit, % (*27*)	33.3 (30.3–35.9)	33.4 (30.7–36.0)
Thrombocytes, G/L (*28*)	210 (169–257)	197 (153–232)
Leukocytes, G/L (*27*)	6.4 (5.2–8.0)	6.4 (5.2–7.8)
Albumin, g/L (*117*)	36.0 (33.0–39.2)	35.0 (32.0–37.5)
C-reactive protein, mg/dL (*96*)	0.66 (0.29–1.9)	0.72 (0.29–2.20)
Antithrombotic therapy		
Anticoagulation therapy	139 (22.2%)	83 (50.3%)
Platelet inhibitor therapy	345 (55.2)	90 (54.6)

Abbreviation: BMI, body mass index.
Note: Italic number in brackets indicates the number of patients with missing values. Continuous variables are presented as median (25th–75th percentile) and categorical variables as absolute frequencies and percentages.
[a]Excluding nonmelanoma skin cancer. Patients with multiple malignancies were counted once.

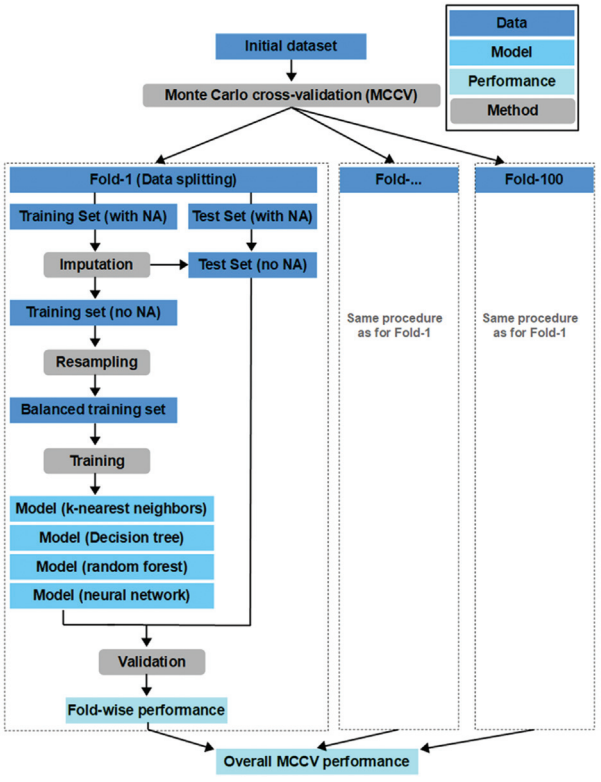

Fig. 1 Machine learning strategy.

using four classification algorithms, including a k-nearest neighbor algorithm, a decision tree algorithm, a random forest algorithm, and a neural network algorithm to derive a total of four classification models. A total of 25 features consisting of patient demographics, clinical characteristics, dialysis-specific parameters, and laboratory parameters were used for the training procedure. For the validation, 100-fold Monte Carlo (MC) cross-validation with 85% of samples in the training set was employed to obtain a highly robust and repeatable estimation of the predictive performance of the model. Preplanned feature selection, feature importance measurements, and hyperparameter tuning were not performed due to the insufficient performance of all models. To identify potential clusters within the data, principal component analysis (PCA) was performed to reduce the dataset to two dimensions. A detailed list of parameters and features used for the analysis can be found in ►**Supplementary Tables S1** and **S2** (available in the online version).

The analysis was performed using Python 3.9.5 and the external packages Scikit-Learn 0.24.2,[33] Seaborn 0.11.1,[34] Matplotlib 3.3.4,[35] Numpy 1.20.2,[36] Pandas 1.2.4,[37] and Imblearn 0.8.0.[38]

Results

Clinical characteristics of all 625 patients are presented in ►**Table 1**. In brief, the median age was 66 (54.5–75.0)

years, 231 (37%) patients were female, and 26% had AF at baseline. During a median follow-up time of 3.47 (3.38–3.58) years, 89 (14.2%) patients developed major bleeding comprising of 41 gastrointestinal bleeds, 14 intracranial hemorrhagic events, 13 other critical organ hemorrhage events, and 21 other major bleeding events. This accounts for a major bleeding rate of 6.8 (95% CI: 5.5–8.4) per 100 patient-years or a cumulative incidence of 6.1% (95% CI: 4.2–8.0) at 1 year, 10.3% (95% CI: 8.0–12.8) at 2 years, and 13.5% (95% CI: 10.8–16.2) at 3 years. One patient (0.16%) was lost to follow-up and 113 (18.1%) received kidney transplantation.

Performance of Bleeding Risk Prediction Tools

Discriminatory performances of seven currently existing bleeding risk prediction scores are shown in ►**Table 2**. In summary, none of the risk assessment models had a C-statistic above 0.60 in the total study population and in the AF population, CIs of all scores overlapped 0.50. Of all scores, the HAS-BLED score[26] showed the highest discriminatory performance in the total population with a C-statistic of 0.59 (95% CI: 0.53–0.66). C-statistics of the ATRIA (Anticoagulation and Risk Factors in Atrial Fibrillation),[27] HEMORR$_2$HAGES (Hepatic or renal disease, Ethanol abuse, Malignancy, Older age, Reduced platelet count or function, Rebleeding risk, Hypertension, Anemia, Genetic factors, Excessive fall risk, and Stroke),[39] ORBIT (Outcomes Registry for Better Informed Treatment),[28] OBRI (Outpatient Bleeding Risk Index),[29] mOBRI (modified Outpatient Bleeding Risk Index),[30] and Shireman et al[31] scores ranged between 0.54 and 0.59. Calibration plots are presented in ►**Fig. S1**. Based upon visual inspection and the Brier score, calibration performance for bleeding risk was moderate across all prediction scores. Detailed comparisons between risk scores reporting on NRI and IDI including decision curve analysis are available in ►**Supplementary Table S3** and ►**Fig. S2** (available in the online version).

Table 2 Discriminatory performance of bleeding risk scores in patients on hemodialysis

Bleeding risk score	Total study cohort ($n = 625$) C-statistics (95% CI)	Subgroup of patients with atrial fibrillation ($n = 165$) C-statistic (95% CI)
HAS-BLED	0.59 (0.53–0.66)	0.54 (0.42–0.66)
ATRIA	0.55 (0.48–0.62)	0.58 (0.46–0.70)
HEMORR$_2$HAGES	0.58 (0.51–0.65)	0.56 (0.44–0.69)
ORBIT	0.59 (0.52–0.66)	0.61 (0.47–0.74)
OBRI	0.54 (0.47–0.61)	0.54 (0.44–0.64)
mOBRI	0.54 (0.47–0.60)	0.55 (0.44–0.65)
Shireman et al	0.59 (0.52–0.67)	0.62 (0.50–0.74)

Abbreviation: CI, confidence interval.
Note: In the total cohort, C-statistics of existing bleeding risk scores ranged between 0.54 and 0.59, indicating poor discriminatory performance. In the subgroup of patients with atrial fibrillation, C-statistics ranged between 0.54 and 0.62, while confidence intervals overlapped 0.5, which indicates that no model performed better than chance.

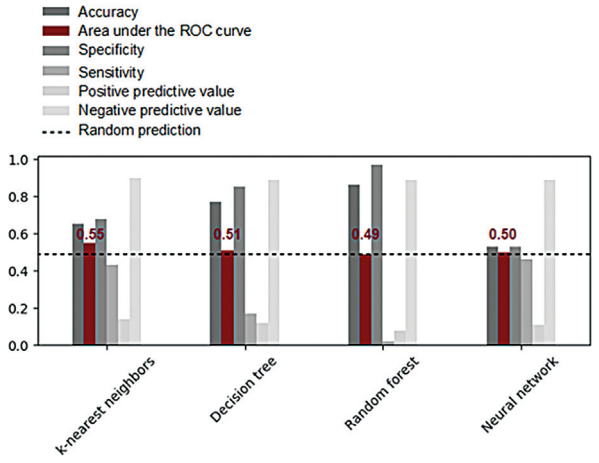

Fig. 2 Performance of four machine learning-based prediction models to predict major bleeding in patients on hemodialysis.

Performance of Machine Learning-Based Bleeding Risk Models

The MC cross-validation procedure yielded area under the receiver operating characteristic curve (AUROC) values close to random predictions for all four machine learning algorithms. AUROC performances were 0.55, 0.51, 0.50, and 0.49 for k-nearest neighbor, decision tree, neural network, and random forest algorithm, respectively (►**Fig. 2**). Accuracies ranged from 0.86 for the random forest to 0.53 for the neural network, indicating imbalanced predictions of the two classes. Sensitivities for decision tree and random forest were low with 0.17 and 0.02, respectively. The other algorithms performed better at handling the imbalanced classes with sensitivities of 0.43 and 0.46 as well as specificities of 0.68 and 0.53 for the k-nearest neighbor algorithm and neural network, respectively. The cross-validation performance metrics, including accuracy, AUROC, sensitivity, specificity, positive predictive value, and negative predictive value for the four algorithms are shown in ►**Supplementary Table S4** (available in the online version). Overall, the performance metrics indicated that none of the machine learning algorithms was able to pick up any patterns associated with the bleeding risk of the patients. These results were further supported by the PCA-based dimensionality reduction, which did not show any cluster formation after reduction to two dimensions based on visual assessment as shown in ►**Fig. 3**.

Discussion

In this prospective study of patients with ESKD on HD, we evaluated currently existing bleeding risk scores, developed in a nondialysis population, for their suitability in this special patient cohort. None of the seven bleeding risk assessment

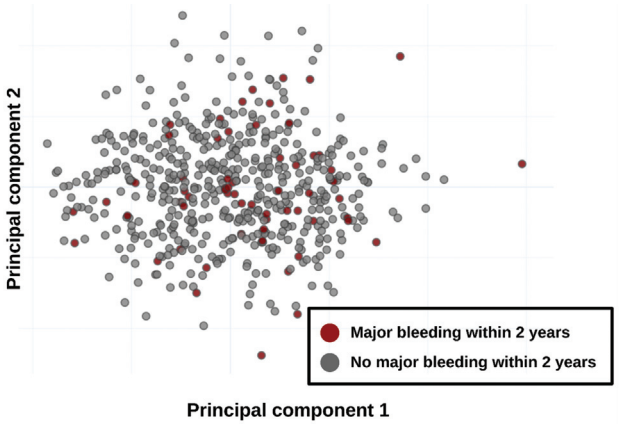

Fig. 3 Principal component analysis in patients on hemodialysis.

tools showed a C-statistic above 0.60, indicating poor performance to identify patients at low or high risk for bleeding. The machine learning-based prediction models derived from and internally validated in our dataset showed similar poor predictive abilities. The four commonly used machine learning models, which were based on 25 features comprising patient demographics, clinical characteristics, and laboratory measurements (shown in ►**Supplementary Table S2**, available in the online version) did not predict major bleeding within 24 months better than chance.

Bleeding risk assessment poses a special challenge to the treating physician when deciding on anticoagulation modalities during HD or anticoagulation for treatment and prevention of thromboembolic events. Therefore, different models have been proposed to assess bleeding risk in clinical practice in the general patient population.[40,41] However, potentially due to the heterogeneity in bleeding type and causality of bleeding events, risk prediction performs only poorly to moderately across different diseases and clinical settings.[24,42] Individual bleeding risk prediction is especially challenging in patient populations at high bleeding risk.[43,44] In our prevalent cohort (i.e., patients who were on HD at enrollment into the study), risk of developing major bleeding was 10% after 2 years. This risk estimate highlights the burden of bleeding in HD patients and is in line with previous findings of similar settings and bleeding definitions.[2,4,8]

Of note, none of the existing bleeding risk assessment strategies withstood external validation in our HD cohort. This is exemplified by the HAS-BLED score, which performed "best" in our cohort. The HAS-BLED score showed a C-statistic of 0.59 in our total cohort of patients with ESKD on HD and 0.54 in the subgroup of patients with AF, indicating poor discrimination of patients who developed major bleeding compared with those who did not. Further, expected event rates differed from observed event rates as shown in ►**Supplementary Fig. S1** (available in the online version). The HAS-BLED score is recommended by several guidelines as the best model to predict bleeding with C-statistics ranging up to 0.80 in some studies.[9–12,45,46] Other bleeding risk scores do not offer advantages over the HAS-BLED score.[47–49] In a randomized controlled trial of AF patients, regular bleeding risk evaluation was associated with a reduction in bleeding events and increased anticoagulation usage.[50] Thus, appropriate and responsible use of bleeding risk scores is suggested to identify high-risk patients and address modifiable bleeding risk factors to improve patient outcome. However, previous studies already highlighted the difficulty of bleeding risk assessment in patients with chronic kidney disease[44] and a prior evaluation of bleeding risk scores (i.e., HAS-BLED, ATRIA, HEMORR$_2$HAGES, ORBIT) in HD patients showed similarly poor predictive results.[51] In our study, we could confirm that these bleeding risk scores perform poorly in ESKD patients on HD. In addition, we have analyzed other bleeding risk scores, such as the OBRI, mOBRI, and the score of Shireman et al, which did not show an improved predictive value. Taken together, all bleeding risk scores performed worse in cohort studies in dialysis patients when compared with their development and validation studies, with most scores not predicting bleeding better than chance. This is true for the overall population of patients with ESKD on HD as well as the subgroup of AF patients and suggests that standard statistical models might not be appropriate for predicting bleeding in this specific patient population.

In need of better prognostic models, we applied machine learning to our dataset. Machine learning represents a major branch of artificial intelligence and data science, and holds promising results in predictive modeling with the ability to reveal key features from complex datasets. Its strength lies in handling enormous numbers of predictors and combining them in nonlinear and highly interactive ways.[52,53] Strikingly, none of four different machine learning-based prediction models, which included 25 features consisting of patient demographics, clinical characteristics, dialysis-specific parameters, and common laboratory parameters, performed better than chance during internal validation in our dataset. Further, the lack of clustering or pattern recognition in PCA confirms the absence of helpful predictors for major bleeding. Earlier studies without machine learning methods concur with our findings while some identified single predictor variables such as recent gastrointestinal bleeding.[54] In summary, no strategy to assess bleeding risk in patients on HD can be recommended to date, as neither standard statistical nor machine-learning methods based on general clinical parameters have proven useful in predicting bleeding events. Thus, novel approaches to characterize bleeding risk in patients on HD are urgently needed. Notably, biomarker-based prediction scores such as the ABC (age, biomarkers, clinical history) bleeding risk score showed promising results in patients with AF and might guide the development of specific bleeding risk scores in patients with ESKD on HD.[55]

Some strengths and limitations must be considered when interpreting our findings. First, the bleeding risk scores evaluated in our study were mainly derived from AF cohorts, and in our population only one-quarter had AF at baseline. Second, the power of machine learning increases by sample size and in contrast to other larger studies that rely on insurance claims databases, our cohort is considered of only moderate size. On the contrary, face-to-face follow-ups and independent adjudication of bleeding events supplied optimal data quality for our

analysis. Third, the ISTH definition for major bleeding was used in our study. Therefore, administration of two or more units of packed erythrocytes was counted as major bleeding. However, patients on HD frequently suffer from renal anemia resulting in a low threshold for blood transfusion and would therefore more likely fulfill the criteria for major bleeding. Nevertheless, fatal bleeding and bleeding in a critical organ including intracranial bleeding contributed substantially to the high incidence of major bleeding in our cohort. Furthermore, we could not consider time-to-event data in our machine learning approach, thus patients who were lost to follow-up had to be excluded from model development. Finally, frequent changes in antithrombotic strategies throughout the study period were noted, which might have affected the risk of bleeding. We tried to address this issue by including a baseline variable reflecting the intention-to-treat approach for antithrombotic medications as a feature in the machine learning models, at the risk of oversimplifying the variables for anticoagulation and antiplatelet therapy.

Conclusion

Bleeding risk assessment in patients with ESKD on HD is suboptimal and remains challenging, as existing prediction tools do not provide useful information for patient management. Similarly, our machine learning approach failed to identify useful models derived from a set of clinically available markers in current practice. Thus, we conclude that common clinical parameters are insufficient to stratify HD patients by bleeding risk and call for the development of novel approaches including biomarker-based prediction scores to characterize the bleeding risk in patients with ESKD on HD.

What is known about this topic?

- Patients with end-stage kidney disease on hemodialysis are at very high risk of bleeding.
- Stratification by bleeding risk could guide clinical decision-making for antithrombotic therapies or dialysis treatment modalities. However, no valid approach toward bleeding risk assessment has been recommended so far.

What does this paper add?

- Our study highlights that existing bleeding risk assessment models—mainly developed in a nondialysis population—are not useful for predicting bleeding in patients with end-stage kidney disease on hemodialysis.
- Also machine learning-based prediction models did not help to identify patients at high risk for bleeding.

Funding
The VIVALDI study was supported by an unrestricted grant from the Austrian National Bank (Jubiläumsfond Österreichische Nationalbank, Project Number 16433), Joseph-Skoda-Prize of the Austrian Association of Internal Medicine (ÖGIM), and the Austrian Science Fund (FWF), Special Research Program (SFB) 54.

Conflict of Interest
None declared.

Acknowledgments
We acknowledge the members of the adjudication committee: Julia Riedl, Johannes Thaler, Christoph Kopp, Thomas Gremmel, and Fritz Leutmezer.

References

1 Lutz J, Menke J, Sollinger D, Schinzel H, Thürmel K. Haemostasis in chronic kidney disease. Nephrol Dial Transplant 2014;29(01): 29–40
2 Sood MM, Bota SE, McArthur E, et al. The three-year incidence of major hemorrhage among older adults initiating chronic dialysis. Can J Kidney Health Dis 2014;1:21
3 Molnar AO, Bota SE, Garg AX, et al. The risk of major hemorrhage with CKD. J Am Soc Nephrol 2016;27(09):2825–2832
4 Wasse H, Gillen DL, Ball AM, et al. Risk factors for upper gastrointestinal bleeding among end-stage renal disease patients. Kidney Int 2003;64(04):1455–1461
5 Trivedi H, Yang J, Szabo A. Gastrointestinal bleeding in patients on long-term dialysis. J Nephrol 2015;28(02):235–243
6 Kuo CC, Kuo HW, Lee IM, Lee CT, Yang CY. The risk of upper gastrointestinal bleeding in patients treated with hemodialysis: a population-based cohort study. BMC Nephrol 2013;14:15
7 Olesen JB, Lip GYH, Kamper A-L, et al. Stroke and bleeding in atrial fibrillation with chronic kidney disease. N Engl J Med 2012;367 (07):625–635
8 Wong CX, Odutayo A, Emdin CA, Kinnear NJ, Sun MT. Meta-analysis of anticoagulation use, stroke, thromboembolism, bleeding, and mortality in patients with atrial fibrillation on dialysis. Am J Cardiol 2016;117(12):1934–1941
9 Hindricks G, Potpara T, Dagres N, et al. 2020 ESC guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. Eur Heart J 2021;42(05):373–498
10 Konstantinides SV, Meyer G, Becattini C, et al; The Task Force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC) 2019 ESC guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS): The Task Force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC). Eur Respir J 2019;54(03):2019
11 Sousa-Uva M, Neumann F-J, Ahlsson A, et al; ESC Scientific Document Group. 2018 ESC/EACTS guidelines on myocardial revascularization. Eur J Cardiothorac Surg 2019;55(01):4–90
12 Meschia JF, Bushnell C, Boden-Albala B, et al; American Heart Association Stroke Council Council on Cardiovascular and Stroke Nursing Council on Clinical Cardiology Council on Functional Genomics and Translational Biology Council on Hypertension. Guidelines for the primary prevention of stroke: a statement for healthcare professionals from the American Heart Association/ American Stroke Association. Stroke 2014;45(12):3754–3832
13 van Eck van der Sluijs A, Abrahams AC, Rookmaaker MB, et al. Bleeding risk of haemodialysis and peritoneal dialysis patients. Nephrol Dial Transplant 2021;36(01):170–175

14 Wang IK, Cheng YK, Lin CL, et al. Comparison of subdural hematoma risk between hemodialysis and peritoneal dialysis patients with ESRD. Clin J Am Soc Nephrol 2015;10(06): 994–1001
15 Huang K-W, Leu H-B, Luo J-C, et al. Different peptic ulcer bleeding risk in chronic kidney disease and end-stage renal disease patients receiving different dialysis. Dig Dis Sci 2014;59(04): 807–813
16 Königsbrügge O, Meisel H, Beyer A, et al. Anticoagulation use and the risk of stroke and major bleeding in patients on hemodialysis: from the VIVALDI, a population-based prospective cohort study. J Thromb Haemost 2021;19(12):2984–2996
17 Van Der Meersch H, De Bacquer D, De Vriese AS. Vitamin K antagonists for stroke prevention in hemodialysis patients with atrial fibrillation: a systematic review and meta-analysis. Am Heart J 2017;184:37–46
18 Hayashi M, Takamatsu I, Kanno Y, Yoshida T, Abe T, Sato YJapanese Calciphylaxis Study Group. A case-control study of calciphylaxis in Japanese end-stage renal disease patients. Nephrol Dial Transplant 2012;27(04):1580–1584
19 Tan J, Liu S, Segal JB, Alexander GC, McAdams-DeMarco M. Warfarin use and stroke, bleeding and mortality risk in patients with end stage renal disease and atrial fibrillation: a systematic review and meta-analysis. BMC Nephrol 2016;17(01):157
20 Königsbrügge O, Lorenz M, Auinger M, et al. Venous thromboembolism and vascular access thrombosis in patients with end-stage renal disease on maintenance hemodialysis: cross-sectional results of the Vienna InVestigation of AtriaL fibrillation and thromboembolism in patients on hemoDIalysis (VIVALDI). Thromb Res 2017;158:59–64
21 Schulman S, Angerås U, Bergqvist D, Eriksson B, Lassen MR, Fisher WSubcommittee on Control of Anticoagulation of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in surgical patients. J Thromb Haemost 2010;8(01):202–204
22 Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. J Clin Epidemiol 2015;68(02):134–143
23 Nopp S, Ay C. Bleeding risk assessment in patients with venous thromboembolism. Hamostaseologie 2021;41(04):267–274
24 Chang G, Xie Q, Ma L, et al. Accuracy of HAS-BLED and other bleeding risk assessment tools in predicting major bleeding events in atrial fibrillation: a network meta-analysis. J Thromb Haemost 2020;18(04):791–801
25 van Es N, Wells PS, Carrier M. Bleeding risk in patients with unprovoked venous thromboembolism: a critical appraisal of clinical prediction scores. Thromb Res 2017;152:52–60
26 Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. Chest 2010;138(05):1093–1100
27 Fang MC, Go AS, Chang Y, et al. A new risk scheme to predict warfarin-associated hemorrhage: the ATRIA (Anticoagulation and Risk Factors in Atrial Fibrillation) Study. J Am Coll Cardiol 2011;58 (04):395–401
28 O'Brien EC, Holmes DN, Thomas L, et al. Therapeutic strategies following major, clinically relevant nonmajor, and nuisance bleeding in atrial fibrillation: findings from ORBIT-AF. J Am Heart Assoc 2018;7(12):e006391
29 Landefeld CS, Goldman L. Major bleeding in outpatients treated with warfarin: incidence and prediction by factors known at the start of outpatient therapy. Am J Med 1989;87(02):144–152
30 Beyth RJ, Quinn LM, Landefeld CS. Prospective evaluation of an index for predicting the risk of major bleeding in outpatients treated with warfarin. Am J Med 1998;105(02):91–99
31 Shireman TI, Mahnken JD, Howard PA, Kresowik TF, Hou Q, Ellerbeck EF. Development of a contemporary bleeding risk model for elderly warfarin recipients. Chest 2006;130(05):1390–1396
32 Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: Synthetic Minority Over-sampling Technique. J Artif Intell Res 2002; 16:321–357
33 Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825–2830
34 Waskom ML. seaborn: statistical data visualization. J Open Source Softw 2021;6(60):3021
35 Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng 2007;9(03):90–95
36 Svd W, Colbert SC, Varoquaux G. The NumPy array: a structure for efficient numerical computation. Comput Sci Eng 2011;13(02): 22–30
37 McKinney W. pandas: a Foundational Python Library for Data Analysis and Statistics. 2011. Accessed February 7, 2022 athttps://www.dlr.de/sc/portaldata/15/resources/dokumente/pyhpc2011/submissions/pyhpc2011_submission_9.pdf
38 Guillaume Lemaître FNaCKA. Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning. J Mach Learn Res 2017;18(17):1–5
39 Gage BF, Yan Y, Milligan PE, et al. Clinical classification schemes for predicting hemorrhage: results from the National Registry of Atrial Fibrillation (NRAF). Am Heart J 2006;151(03):713–719
40 Nopp S, Ay C. Bleeding risk assessment in patients with venous thromboembolism. Hamostaseologie 2021;41(04):267–274
41 Zulkifly H, Lip GYH, Lane DA. Bleeding risk scores in atrial fibrillation and venous thromboembolism. Am J Cardiol 2017; 120(07):1139–1145
42 Riva N, Bellesini M, Di Minno MN, et al. Poor predictive value of contemporary bleeding risk scores during long-term treatment of venous thromboembolism. A multicentre retrospective cohort study. Thromb Haemost 2014;112(03):511–521
43 de Winter MA, Dorresteijn JAN, Ageno W, et al. Estimating bleeding risk in patients with cancer-associated thrombosis: evaluation of existing risk scores and development of a new risk score. Thromb Haemost 2021. Doi: 10.1055/s-0041-1735251
44 McAlister FA, Wiebe N, Jun M, et al. Are existing risk scores for nonvalvular atrial fibrillation useful for prediction or risk adjustment in patients with chronic kidney disease? Can J Cardiol 2017; 33(02):243–252
45 Borre ED, Goode A, Raitz G, et al. Predicting thromboembolic and bleeding event risk in patients with non-valvular atrial fibrillation: a systematic review. Thromb Haemost 2018;118(12):2171–2187
46 Chao T-F, Joung B, Takahashi Y, et al. 2021 focused update consensus guidelines of the Asia Pacific Heart Rhythm Society on Stroke Prevention in Atrial Fibrillation: executive summary. Thromb Haemost 2022;122(01):20–47
47 Proietti M, Romiti GF, Vitolo M, Potpara TS, Boriani G, Lip GYH. Comparison of HAS-BLED and ORBIT bleeding risk scores in AF patients treated with NOACs: a report from the ESC-EHRA EORP-AF general long-term registry. Eur Heart J Qual Care Clin Outcomes 2021. Doi: 10.1093/ehjqcco/qcab069
48 Wattanaruengchai P, Nathisuwan S, Karaketklang K, Wongcharoen W, Phrommintikul A, Lip GYH. Comparison of the HAS-BLED versus ORBIT scores in predicting major bleeding among Asians receiving direct-acting oral anticoagulants. Br J Clin Pharmacol 2022;88 (05):2203–2212
49 Gao X, Cai X, Yang Y, Zhou Y, Zhu W. Diagnostic accuracy of the HAS-BLED bleeding score in VKA- or DOAC-treated patients with atrial fibrillation: a systematic review and meta-analysis. Front Cardiovasc Med 2021;8(1543):757087
50 Guo Y, Lane DA, Chen Y, Lip GYHmAF-App II Trial investigators. Regular bleeding risk assessment associated with reduction in bleeding outcomes: the mAFA-II randomized trial. Am J Med 2020;133(10):1195.e21202.e2

51 Ocak G, Ramspek C, Rookmaaker MB, et al. Performance of bleeding risk scores in dialysis patients. Nephrol Dial Transplant 2019;34(07):1223–1231
52 Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2014;13:8–17
53 Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med 2016;375(13): 1216–1219
54 Sood MM, Larkina M, Thumma JR, et al. Major bleeding events and risk stratification of antithrombotic agents in hemodialysis: results from the DOPPS. Kidney Int 2013;84(03): 600–608
55 Hijazi Z, Oldgren J, Lindbäck J, et al; ARISTOTLE and RE-LY Investigators. The novel biomarker-based ABC (age, biomarkers, clinical history)-bleeding risk score for patients with atrial fibrillation: a derivation and validation study. Lancet 2016;387 (10035):2302–2311