Clinical Kidney Journal, 2024, vol. 17, no. 12, sfae347

https://doi.org/10.1093/ckj/sfae347
Advance Access Publication Date: 15 November 2024
Original Article

ORIGINAL ARTICLE

Prediction models for ischemic stroke and bleeding in dialysis patients: a systematic review and meta-analysis

Christoforos K. Travlos[1,2,3], Adario Chirgwin-Dasgupta[4], Emilie Trinh[1,2,3], Allan D. Sniderman[5], Ahsan Alam[1,2,3] and Thomas A. Mavrakanas[1,2,3]

[1]Division of Nephrology, Department of Medicine, McGill University Health Centre, Montreal, Quebec, Canada, [2]Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada, [3]Division of Clinical and Translational Research, Department of Medicine, McGill University, Montreal, Quebec, Canada, [4]Division of General Internal Medicine, Department of Medicine, McGill University Health Centre, Montreal, Quebec, Canada and [5]Mike and Valeria Rosenbloom Centre for Cardiovascular Prevention, Department of Medicine, McGill University Health Centre, Montreal, Quebec, Canada

Correspondence to: Christoforos K. Travlos; E-mail: christoforos.travlos@mail.mcgill.ca

Watch the video abstract of this contribution

ABSTRACT

Background. Patients with kidney failure on maintenance dialysis have a high stroke and bleeding risk. Multivariable prediction models can be used to estimate the risk of ischemic stroke and bleeding. A systematic review and meta-analysis was performed to determine the performance of the existing models in patients on dialysis.
Methods. MEDLINE and Embase databases were searched, from inception through 12 January 2024, for studies of prediction models for stroke or bleeding, derived or validated in dialysis cohorts. Discrimination measures for models with c-statistic data from three or more cohorts were pooled by random effects meta-analysis and a 95% prediction interval (PI) was calculated. Risk of bias was assessed using PROBAST. The review was conducted according to the PRISMA statement and the CHARMS checklist.
Results. Eight studies were included in this systematic review. All the included studies validated pre-existing models that were derived in cohorts from the general population. None of the identified studies reported the development of a new dialysis specific prediction model for stroke, while dialysis specific risk scores for bleeding were proposed by two studies. In meta-analysis of c-statistics, the CHA_2DS_2-VASc, $CHADS_2$, ATRIA, HEMORR(2)HAGES and HAS-BLED scores showed very poor discriminative ability in the dialysis population. Six of the eight included studies were at low or unclear risk of bias and certainty of evidence was moderate.
Conclusions. The existing prediction models for stroke and bleeding have very poor performance in the dialysis population. New dialysis-specific risk scores should be developed to guide clinical decision making in these patients.

Keywords: bleeding, dialysis, meta-analysis, prediction models, stroke

Received: 17.10.2024; Editorial decision: 12.11.2024

KEY LEARNING POINTS

What was known:

- Patients with end-stage kidney disease on maintenance dialysis have a high risk of stroke and bleeding.
- In the general population, risk scores that predict the risk of stroke or bleeding are used to guide clinical decisions.
- These scores were developed in cohorts that did not include patients on dialysis and their performance when used in this population has never been studied in a systematic way.

This study adds:

- This systematic review of current literature revealed no existing stroke prediction models specifically designed for dialysis patients.
- Only two studies have developed bleeding risk scores for patients on dialysis, but their performance is poor.
- The most widely used, in the general population, stroke and bleeding risk scores like CHA_2DS_2-VASc and HAS-BLED have very poor performance in patients on dialysis, making risk stratification impossible.

Potential impact:

- The development and validation of new stroke and bleeding risk scores specifically designed for the dialysis population must be a high priority.

INTRODUCTION

Stroke is the second leading cause of death and disability worldwide [1]. Although stroke mortality rates have declined over the last two decades, the incidence and the burden of the disease have increased [2]. Chronic kidney disease (CKD) constitutes an independent risk factor for stroke [3, 4], while patients with end-stage kidney disease (ESKD) on renal replacement therapy have an overall risk up to 10 times higher than the general population [5, 6]. In addition, the prognosis of patients with ESKD suffering from stroke is much poorer in comparison with other populations [7].

A combination of common cardiovascular risk factors that are highly prevalent in the dialysis population, such as older age, atrial fibrillation (AF), diabetes and dyslipidemia, alongside unique characteristics related to dialysis (hemodynamic changes and variability in blood pressure, vascular calcification, chronic inflammation and uremic platelet dysfunction) compromise hemostatic mechanisms. As a result, not only is an increased risk of stroke observed in patients on dialysis, but also a higher number of major bleeding episodes.

In the general population, risk scores that predict the risk of stroke or hemorrhage are used to guide clinical decisions. The majority of them have been developed and widely validated in cohorts of patients with AF. The most used risk scores for the prediction of stroke are the $CHADS_2$ (Congestive heart failure, Hypertension, Age ≥75 years, Diabetes Mellitus, prior Stroke) [8] and CHA_2DS_2-VASc (Congestive heart failure, Hypertension, Age ≥75 years, Diabetes Mellitus, prior Stroke, Vascular disease, Age 65–74 years, female Sex) [9]. Both have demonstrated high discriminative performance in the general population with a c-statistic of 0.812 [95% confidence interval (CI) 0.796–0.827] for the $CHADS_2$ and 0.888 (95% CI 0.875–0.900) for the CHA_2DS_2-VASc [10]. Other less commonly used scores include the R_2CHADS_2 (Renal dysfunction, Congestive heart failure, Hypertension, Age ≥75 years, Diabetes Mellitus, prior Stroke), the ATRIA score, [11] the Q-stroke algorithm and the GARFIELD-AF score. For bleeding risk prediction the HAS-BLED [12], the HEMORR(2)HAGES [13] and the ATRIA [11] are the most widely used scores.

All the aforementioned prediction models were developed in AF cohorts that did not include patients on dialysis. Thus, their simple extrapolation in this population—that is characterized by an elevated baseline stroke and bleeding risk, independently of AF status—is a matter of debate.

It is not clearly established whether the existing scores perform adequately in dialysis patients and there are no widely used risk scores specifically designed for this population. Therefore, the aim of this review is to summarize in a systematic way all the data regarding the development of new or validation of already existing predictive scores for stroke and bleeding in dialysis patients with or without AF.

MATERIALS AND METHODS

Search strategy

MEDLINE and Embase databases were systematically searched through the Ovid platform from inception through 12 January 2024. A combination of keywords and subject headings related to stroke, bleeding, ESKD and prediction models was used (online Supplementary data). Forward and backward citation searching for included studies and previous systematic reviews was performed. Endnote's duplicate detection function was used to detect duplicates.

Study selection

Studies were eligible if they met the following inclusion criteria: (i) original study in adults (≥18 years of age); (ii) study that reported the development or validation of prediction models for stroke or bleeding; (iii) the development or validation cohort must report results for patients with ESKD; and (iv) articles should be published in peer-reviewed journals.

We excluded: (i) studies reporting the association between risk factors with stroke and bleeding and not a full prediction model; and (ii) studies with prediction models for stroke or bleeding in the general population. To be included in the meta-analysis, a model had to have c-statistic data from three or more cohorts.

Records were uploaded to a systematic review web application (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia). Two independent investigators (C.K.T., A.C.-D.) screened the titles and abstracts of all identified articles and, when required, reviewed full-text manuscripts to

identify relevant studies for inclusion. Disagreements were resolved by discussion and if consensus could not be reached a third senior investigator (T.A.M.) was consulted.

This review was registered on PROSPERO (CRD42024543041) and was conducted according to the PRISMA statement and CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS; online Supplementary data) [14, 15].

Data extraction and quality assessment

The full text and supplement of included studies was reviewed by two independent investigators (C.K.T., A.C.-D.) who extracted data in an excel spreadsheet using the same protocol. The same two investigators assessed each model in each study for risk of bias and applicability to the review question using the Prediction model Risk Of Bias ASsessment Tool (PROBAST) [16]. Discrepancies between reviewers were resolved with consensus.

The Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach was used to assess the certainty of the evidence [17].

Measures of discrimination and calibration were extracted, to allow the quantitative synthesis of the models' predictive performance [18]. We extracted data on the c-statistic or area under the receiver operating characteristic curve (AUROC) and corresponding 95% CI. When the 95% CI was not reported, we calculated it using the methods described by Debray *et al.* [18]. We extracted data on whether the ratio for observed to expected (O:E) events or calibration slope were reported. When augmentation of pre-existing models was performed by adding variables in order to enhance the predictive value of models, we retrieved the rate of improvement of the net reclassification improvement (NRI) index of the augmented model compared with the original model as well as the *P*-value for the difference between the two models.

Statistical analysis

Continuous variables were reported as means ± standard deviation and categorical variables as percentages. In case that results were reported as median and interquartile range, we converted them in sample mean and standard deviation by using the Wan *et al.* method [19]. A *P*-value $<.05$ was considered statistically significant.

In each study, the c-statistic/AUROC of the model was assessed, where a 95% CI containing 0.5 was indicative of insufficient discrimination. Calibration of the prediction models was assessed according to the method that was reported in each study. Regarding augmentation, we defined significant improvement as an NRI index with a *P*-value of $<.05$ [20].

We conducted a random effects meta-analysis of discrimination through a summary measure of c-statistic and corresponding 95% CI as suggested by Debray *et al.* [18] We calculated the 95% prediction interval (PI) to demonstrate the extent of between-study heterogeneity and to indicate a possible range for prediction model performance in a new validation [21]. When the 95% CI or PI of the summary c-statistic included 0.5, we considered that there was insufficient evidence that the prediction model has statistically significant discriminatory ability for stroke or bleeding in the dialysis population [22, 23]. Summary c-statistics of <0.60, 0.60–0.70, 0.70–0.80 and >0.80 were defined as inadequate, adequate, acceptable and excellent based on previous literature [24].

In our analysis, we assessed overall discrimination for models that had three or more cohorts with c-statistic data. We performed sensitivity analyses restricting our analyses to only those studies where the overall PROBAST assessment was "low" or "unclear" risk of bias. All statistical analyses were performed using Stata (version 18 SE; College Station, TX, USA).

RESULTS

Study selection

A total of 5917 studies were retrieved (2945 from MEDLINE and 2972 from Embase). After duplicates removal 5129 studies remained for screening and 16 studies were included for full text review. Finally, eight studies were included in the systematic review and meta-analysis [25–32]. Full study selection process is shown in Fig. 1.

Characteristics of included studies

The included studies were based on seven different cohorts. Study characteristics and baseline characteristics of the patients in the included cohorts are shown in Table 1 and Supplementary data, Table S1 [25–32]. The number of participants ranged from 141 to 53 147, mean age 60–77 years, proportion of women 34–53%, and mean follow-up of 16 months to 3.5 years.

In three out of five studies that reported the validation of prediction models for ischemic stroke, the authors assessed the performance of the models in patients with ESKD that either had AF before initiation of dialysis therapy or developed AF after dialysis initiation. In the study by Sab *et al.* [29] the AUROC curve and the c-statistic were available only in the total cohort including patients with and without AF. In the remaining one study, de Jong *et al.* [31] validated several models in patients on dialysis, without accounting for the presence or absence of AF. However, they performed sensitivity analysis validating the scores only in patients on dialysis that were concomitant vitamin K antagonist (VKA) users.

In the studies that reported validation of bleeding prediction models, these were tested in cohorts of patients with and without AF. In two studies [26, 31] sensitivity analysis was performed according to VKA use status.

Characteristics of included prediction models

Prediction models for ischemic stroke

In terms of ischemic stroke prediction, all the included studies validated pre-existing models that were derived in cohorts from the general population. None of the identified studies reported the development of a new dialysis specific prediction model for stroke.

Supplementary data, Table S2 summarizes the predictor variables used in each of the included models.

Prediction models for bleeding

Two studies [25, 32] developed new dialysis specific prediction models for bleeding. Madken *et al.* reported the development of the BLEED-HD risk algorithm which was derived in the Dialysis Outcomes and Practice Patterns Study (DOPPS) [33] cohort, using Cox proportional hazards regression. The model was subsequently validated in the Canadian Organ Replacement Registry.

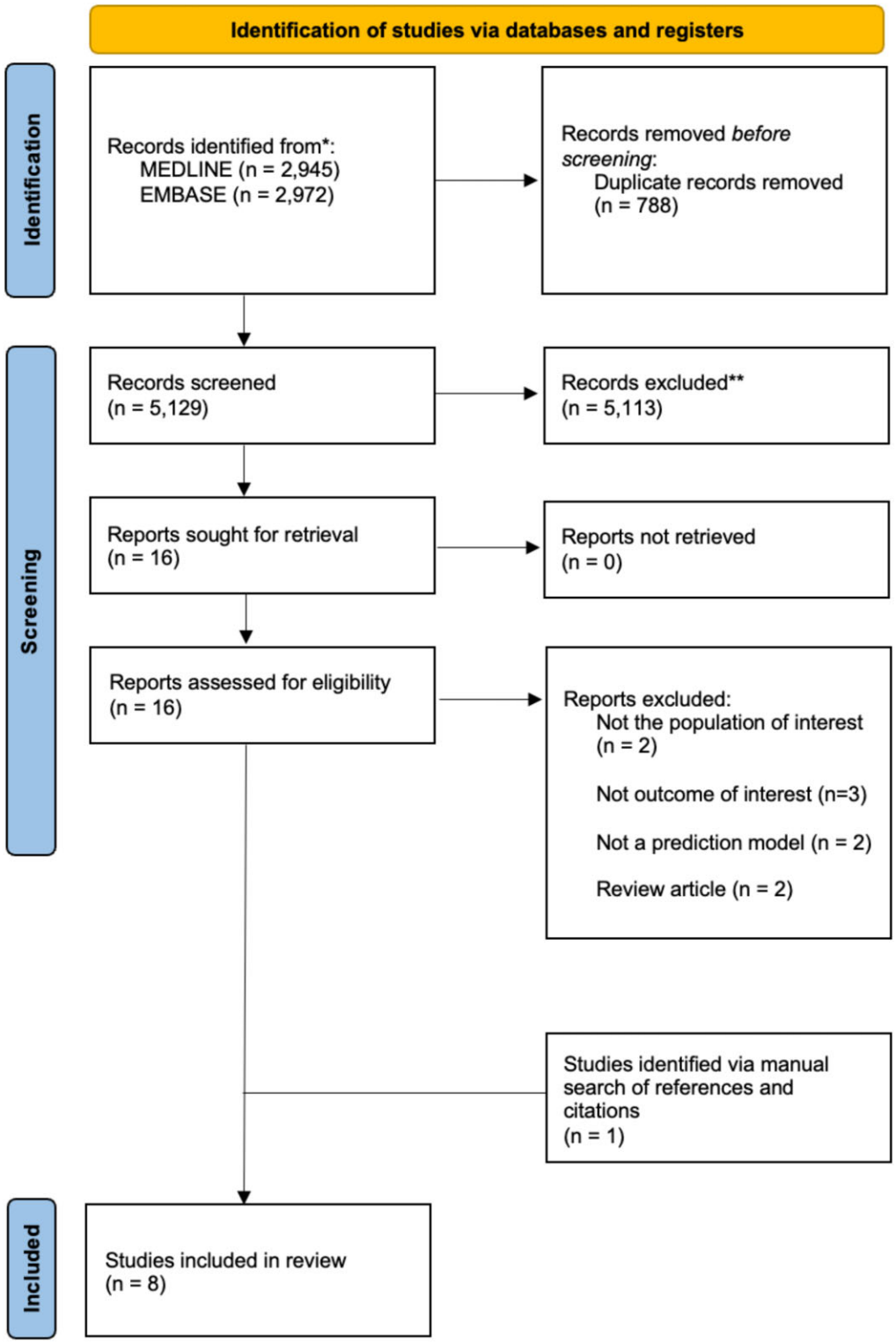

Figure 1: The Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) flow-chart of studies included in the systematic review.

Nopp *et al.* [25] developed four models using machine learning techniques, without externally validating them.

Regarding validation of existing bleeding prediction models, all the included studies validated in dialysis cohorts bleeding prediction algorithms that are used in the general population [11–13, 34, 35].

Supplementary data, Table S2 summarizes the predictor variables used in each of the included models.

Table 1: Characteristics of included studies.

Study	Cohort (country)	Study aim	Outcome	Events/total patients (%)	Age (years, mean ± SD)	Female sex (%)	AF (%)	DM (%)	HTN (%)	Prior stroke (%)	OAC users (%)	CAD (%)	Outcome coding	Enrolment period (mean follow-up in years)
Chao 2014	NHIRD (Taiwan)	EV	Ischemic stroke	1217/10 999 (11.7)	71.0 (11.1)	53.8	100	56.7	90.7	43.2	0	NR	ICD-9-CM	1996–2011 (1.4)
Shih 2016	NHIRD (Taiwan)	EV	Ischemic stroke	600/6772 (8.9)	68.8 (11.3)	53.2	100	58.4	64.9	33.0	8.4	33.0	ICD-9-CM	1998–2011 (3.2)
Sab 2022	Two dialysis facilities (Lebanon)	EV	Ischemic stroke or TIA	22/256 (8.6)	69.3 (13.9)	34.0	100	51.2	96.9	NR	2.7	17.0	EHR	2010–2019 (3.0)
			Bleeding	84/256 (33)										
Wang 2016	Middlemore Hospital (New Zealand)	EV	Ischemic stroke	15/141 (10.6)	61.2 (11.3)	38.0	100	59.6	92.9	20.6	41.8	71.6	EHR	2000–2008 (3.4)
			Bleeding	41/141 (32.6)										
de Jong 2020	NECOSAD (The Netherlands)	EV	Ischemic stroke	127/1955 (6.5)	60.0 (15.1)	37.8	NR	19.8	55.8	7.5	11.3	9.9	NR	1997–2007 (2.5 median)
Ocak 2019	NECOSAD (The Netherlands)	EV	Bleeding	183/1745 (10.4)	61.0 (16.0)	39.0	NR	NR	21.0	8.0	13	NR	ERA codes	1997–2007 (3.0)
Nopp 2022	VIVALDI (Austria)	D, EV	Bleeding	89/625 (14.2)	65.0 (15.2)	37.0	26.4	38	91.8	21.9	22.2	37.1	NR	2014–2015 (3.5)
Madken 2023	DOPPS (multinational)	D, EV	Bleeding	2770/53 147 (5.2)	63.7 (14.7)	39.7	12.3	40.6	84.5	16.0	30.3	40.5	NR	2002–2018 (1.5)
	CORR (Canada)			2406/19 318 (12.5)	76.7 (6.5)	39.2	24.6	63.9	95.1	10.1	17.1	58.7		2008–2019 (1.95)

SD, standard deviation; CAD, coronary artery disease; CORR, Canadian Organ Replacement Registry; D, development; DM, diabetes mellitus; DOPPS, Dialysis Outcomes and Practice Patterns Study; ERA, European Renal Association; EV, external validation; HTN, hypertension; ICD, International Classification of Diseases; NECOSAD, Netherlands Cooperative Study on the Adequacy of Dialysis; NHIRD, National Health Insurance Research Database; TIA, transient ischemic attack; VIVALDI, Vienna InVestigation of Atrial fibrillation and thromboembolism in patients on hemoDialysIs.

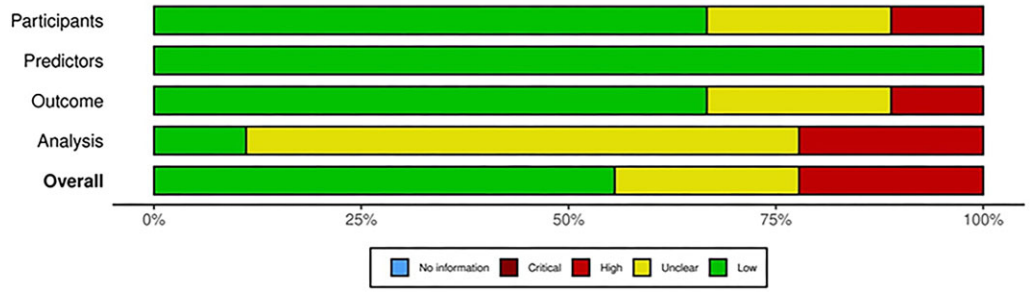

Figure 2: Risk of bias across all included studies.

Risk prediction model performance among included cohorts

Prediction models for ischemic stroke

The discriminative ability of the $CHADS_2$ and CHA_2DS_2-VASc scores that were assessed in three and five studies accordingly was in general poor. In all but one of the studies, the reported c-statistics/AUROC ranged from 0.61 (95% CI 0.60–0.62) [27] to 0.68 (95% CI 0.67–0.69). Only Wang *et al.* [30] found excellent discriminative ability for both models, with a c-statistic of 0.88 (95% CI 0.80–0.96) for the $CHADS_2$ score and 0.85 (95% CI 0.77–0.93) for the CHA_2DS_2-VASc. de Jong *et al.* assessed another 13 models that showed very poor performance (Supplementary data, Table S1).

Calibration of the models was assessed only by de Jong *et al.*, who reported poor calibration of all the fifteen models (including $CHADS_2$ and CHA_2DS_2-VASc) apart from the Framingham Heart Study model [36].

Prediction models for bleeding

Prediction models derived in dialysis cohorts. The BLEED-HD model [32] showed poor discrimination in its developmental cohort with c-statistic ranging from 0.66 to 0.69 over years 1 to 3, but it was well calibrated (observed versus predicted risk difference <2%). Similarly, the four machine learning models developed by Nopp *et al.* [25] had very poor discriminative ability, with c-statistics ranging from 0.49 to 0.55.

External validation of models in dialysis cohorts. The HAS-BLED score was the most frequently assessed bleeding risk score (five studies) with a discriminative ability ranging from 0.50 (95% CI 0.39–0.60) [30] to 0.76 (95% CI 0.69–0.83) [29].

The ATRIA and HEMORR(2)AGES scores were assessed in three studies with c-statistics from 0.55 (95% CI 0.48–0.62) to 0.58 (95% CI 0.51–0.65). The ORBIT, OBRI and mOBRI also showed very poor discriminative performance.

Calibration of the bleeding risk scores was described as poor to moderate in the studies that reported calibration measures.

Augmentation of included risk prediction models

Augmentation data were available in two studies. Chao *et al.* [27] reported that the CHA_2DS_2-VASc score improved the NRI by 4.8% (95% CI 3.2%–6.6%) compared with the $CHADS_2$ score ($P < .0001$), while BLEED-HD score showed an improved NRI over HAS-BLED, ATRIA and HEMORR(2)HAGES scores ($P < .0001$).

Risk of bias assessment

Overall, six of the included studies were at low or unclear risk of bias, while two of them were at high risk of bias (Fig. 2, Supplementary data, Fig. S1).

Meta-analysis

Prediction models for ischemic stroke

Only the CHA_2DS_2-VASc and $CHADS_2$ scores were eligible for meta-analysis. Despite high heterogeneity, the CHA_2DS_2-VASc had a statistically significant 95% PI. Although the prediction interval did not include 0.5, its discriminative performance was far from optimal with a pooled c-statistic of 0.68 (deemed adequate as per our definition). The $CHADS_2$ score had non-significant 95% PI and showed similar discriminative ability as the CHA_2DS_2-VASc score (Fig. 3).

In our sensitivity analysis including only studies not at high risk of bias, which was feasible only for the CHA_2DS_2-VASc score, the summary predictive ability remained adequate (c-statistic of 0.65) but the 95% PI was very large (0.02–1.28) (Supplementary data, Fig. S2a). The 95% prediction interval indicates a possible range for model performance in a new validation. In this case, with values well below 0.5, there was insufficient evidence to suggest that the prediction model has statistically significant discriminatory ability for stroke in the dialysis population.

Prediction models for bleeding

Three bleeding risk scores were eligible for meta-analysis. None of the scores had a 95% PI that did not include 0.5, while ATRIA and HEMORR(2)HAGES showed inadequate predictive performance with no heterogeneity (Fig. 4). The predictive ability of HAS-BLED was borderline adequate but in the sensitivity analysis performed, including only studies not at high risk of bias, the summary c-statistic was 0.56 (suggestive of inadequate discriminative ability) and the 95% PI included 0.5 (Fig. 4, Supplementary data, Fig. S2b).

Certainty of evidence

The overall certainty level was downgraded to “moderate,” because of inconsistent results given high heterogeneity in included studies. This implied that further research is likely to have an important impact on our confidence in the effect estimate.

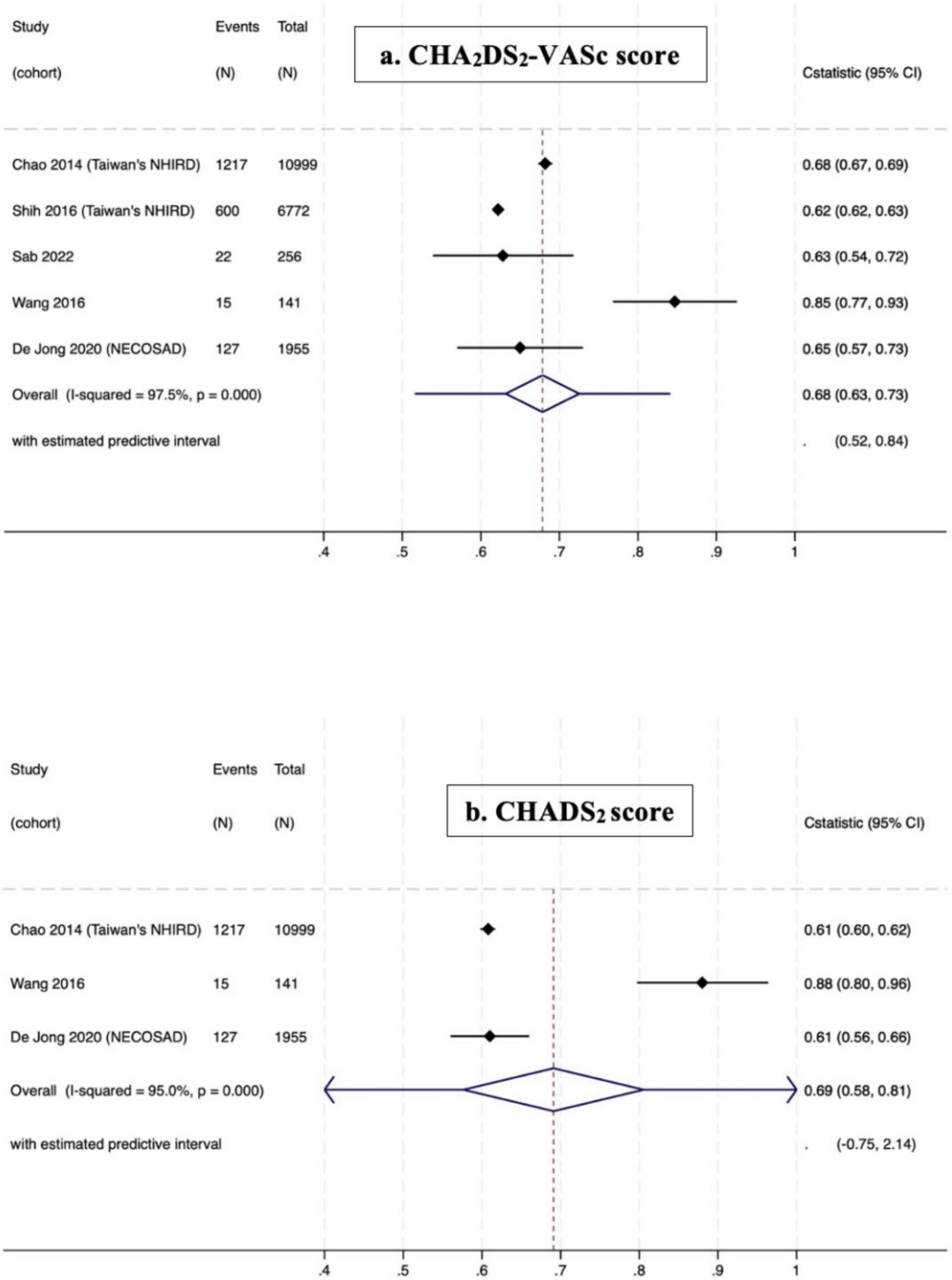

Figure 3: Meta-analysis of c-statistics for: (a) CHA_2DS_2-VASc score and (b) $CHADS_2$ score.

DISCUSSION

This systematic review provides an overview of the studies that reported the development or validation of predictive models for stroke and bleeding in patients on dialysis. The predictive performance of the already existing predictive models that were not developed to be specific for dialysis patients, was poor for both the outcomes of stroke and bleeding. Only Wang *et al*. [30] found excellent discriminative ability for CHA_2DS_2-VASc and $CHADS_2$ scores but given the very small number of participants (141 patients) and outcome events, this study was considered to be at high risk of bias as the reported results may be due to chance or model overfitting. None of the identified studies reported the development of a new, dialysis-specific predictive model for stroke, while in the two studies that derived specific models for bleeding, these did not achieve adequate performance levels.

Patients with ESKD experience significantly elevated risks for ischemic stroke and major hemorrhage—3 to 10 times higher for ischemic stroke and up to 5 times higher for major bleeding—compared with the general population [37–39]. However, the epidemiology of stroke seems to be different in dialysis patients

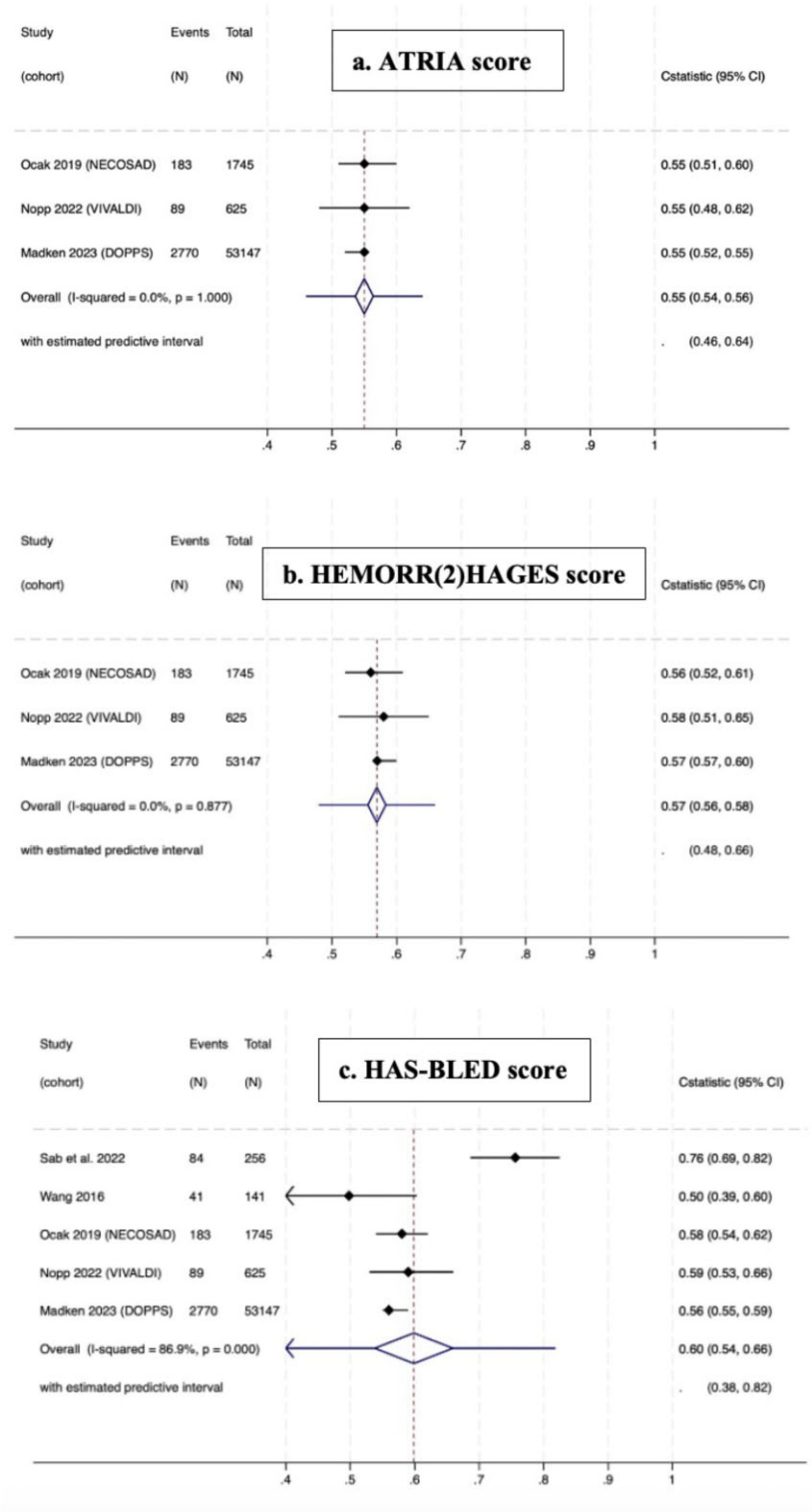

Figure 4: Meta-analysis of c-statistics for: (**a**) ATRIA score; (**b**) HEMORR(2)HAGES score and (**c**) HAS-BLED score.

compared with patients with less advanced CKD or the general population. A recent systematic review by Zamberg *et al.* highlights the fact that the frequency of hemorrhagic stroke is higher in ESKD and approaches the frequency of ischemic stroke [40].

It is obvious that dialysis patients constitute a very specific population, with high prevalence of AF and risk for stroke and bleeding. Thus, appropriate risk stratification is of major importance for their management [41].

Two observational studies have shown that the risk of stroke increases with rising $CHADS_2$ scores [33, 42]. However, in our analysis we showed that the stroke risk scores which are used in the general population to inform clinical decision making and

standardize the use of oral anticoagulation (OAC) may be inappropriate for use in ESKD. This could be explained by the very high prevalence of the classic cardiometabolic variables used in most of the existing scores (age, diabetes, hypertension, vascular disease etc.) in the dialysis population. Consequently, the vast majority of these patients would be classified as of very high risk according to the existing models. Considering also that the incidence of stroke seems to be lower in dialysis patients compared with no-dialysis patients with similar CHA_2DS_2-VASc scores, we could say that CHA_2DS_2-VASc score overestimates stroke risk in the dialysis population patients [43].

The only way to tackle this problem and achieve appropriate risk stratification is by the development of dialysis-specific risk scores. We presented in our analysis that two studies reported bleeding scores for patients with ESKD although none of them had acceptable performance when validated. Nevertheless, no dialysis specific multivariable model for stroke has been published to date. Only De Vriese and Heine have proposed an algorithm to predict the net benefit from OAC. The "Dialysis Risk Score" is based on the CHA_2DS_2-VASc score and is calculated by considering a history of ischemic stroke/transient ischemic attack (3 points), age >75 years (1 point) a history of diabetes (1 point), and a history of clinically relevant gastrointestinal bleeding (–1 point). According to the authors, applying a threshold of ≥2 points, only 44% of Valkyrie trial participants had an indication for OAC [44]. This score was not included in our review as the formal analysis behind its development was not published.

So far, AF is conventionally considered as a binary entity (present or absent) and its existence constitutes a risk factor for stroke. However, due to the high heterogeneity within the AF population, recent studies suggest that the AF burden, which refers to the time spent in AF, may play a more important role when it comes to risk stratification [45].

In a retrospective study of 21 768 nonanticoagulated patients with Cardiac Implantable Electronic Devices (CIED), Kaplan *et al*. examined the utility of AF burden in the determination of stroke risk [46]. They found that stroke risk crossed an actionable threshold, defined as >1%/year in patients with a CHA_2DS_2 -VASc score of 2 with >23.5 h of daily AF duration and those with a CHA_2DS_2-VASc score of 3 to 4 with >6 min of AF duration. In another study, Boriani *et al*. also concluded that the addition of burden increased the ability of CHA_2DS_2-VASc score to discriminate risk for thromboembolism [47]. It would be very plausible to think that the interaction between clinical risk factors and AF burden which would be translated to a CHA_2DS_2-VASc-AFBurden score may inform stroke risk [48]. Such a score might be of particular interest for the appropriate prediction of stroke in the dialysis patients that have high CHA_2DS_2-VASc scores.

It is true that the relationship between cardiac structural changes, atrial remodeling, thrombogenesis vascular risk factors and proarrhythmic milieu in ESKD is very complex. Also, the volume overload, fluid and electrolytes shift, and drastic hemodynamic changes create a very specific disease environment. Thus, factors that specifically pertain to dialysis population like volume status, electrolyte changes and dialysis prescription would be very important variables in the development of new dialysis specific risk scores for the prediction of stroke.

Uremic platelet dysfunction, malnutrition, anemia and endothelial dysfunction all contribute to the high hemorrhagic diatheses of dialysis patients. In our analysis we found that the commonly used bleeding risk scores had very poor performance when tested in dialysis patients. Interestingly, neither of the two studies that developed bleeding scores specifically for this population succeeded.

The strengths of this analysis are a comprehensive search strategy, a thorough analytical approach and an exclusive focus on the dialysis population. However, we acknowledge limitations in our study. We did not investigate between study heterogeneity applying meta-regression or subgroup meta-analysis. Also, almost half of the included studies were at high or unclear risk of bias, although this is very common in predictive research. Meta-analysis of model calibration performance was not possible given the lack of reporting of such data. Finally, patients with and without oral anticoagulation were analyzed together in some cohorts.

The lack of a specific tool for stroke and bleeding risk in ESKD makes risk stratification impossible. The result is conflicting guidelines that breed variable clinical practice patterns, as illustrated by the Dialysis Outcomes and Practice Patterns Study data, published by Wizemann *et al*. [42]. The consequence is that many patients who will not benefit from OAC receive them, whereas others who might, do not. Given that the side effects of OAC are potentially fatal, such imprecision in decision-making should not be acceptable.

Accordingly, developing and validating new predictive tools for stroke and bleeding that are specific for the dialysis population must be a high priority.

SUPPLEMENTARY DATA

Supplementary data are available at *Clinical Kidney Journal* online.

FUNDING

None.

AUTHORS' CONTRIBUTIONS

T.A.M. developed the concept and provided guidance; C.K.T. collected and assembled data, carried out data analysis and interpretation, risk of bias assessment and wrote the initial draft of the manuscript; A.C.-D. collected data, carried out risk of bias assessment; E.T., A.D.S. and A.A. provided guidance; all authors reviewed and approved the final draft of the manuscript.

DATA AVAILABILITY STATEMENT

The data underlying this article are available in the article and in its online supplementary material.

CONFLICT OF INTEREST STATEMENT

T.A.M. has received speaker honoraria from Bayer, BMS Canada, Janssen, AstraZeneca and Pfizer; and has served on advisory boards for Boehringer Ingelheim, Bayer, GSK and Servier outside the submitted work. He has also received research grants from AstraZeneca and Pfizer. He is supported by a Fonds de Recherche Santé Quebec (FRSQ) Junior 1 Clinician Scholar award and a Kidney Research Scientist Core Education and National Training (KRESCENT) program New Investigator Award. E.T. has received speaker honorarium from Baxter Inc., investigator-initiated funding from Otsuka and GSK, and consulting fees from Otsuka. The remaining authors have no relevant disclosures.

REFERENCES

1. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2019;**18**:439–58. https://doi.org/10.1016/S1474-4422(19)30034-1
2. Feigin VL, Norrving B, Mensah GA. Global burden of stroke. *Circ Res* 2017;**120**:439–48. https://doi.org/10.1161/CIRCRESAHA.116.308413
3. Toyoda K, Ninomiya T. Stroke and cerebrovascular diseases in patients with chronic kidney disease. *Lancet Neurol* 2014;**13**:823–33. https://doi.org/10.1016/S1474-4422(14)70026-2
4. Krishnan AV, Kiernan MC. Neurological complications of chronic kidney disease. *Nat Rev Neurol* 2009;**5**:542–51. https://doi.org/10.1038/nrneurol.2009.138
5. Murray AM, Seliger S, Lakshminarayan K *et al*. Incidence of stroke before and after dialysis initiation in older patients. *J Am Soc Nephrol* 2013;**24**:1166–73. https://doi.org/10.1681/ASN.2012080841
6. Arnold J, Sims D, Ferro CJ. Modulation of stroke risk in chronic kidney disease. *Clin Kidney J* 2016;**9**:29–38. https://doi.org/10.1093/ckj/sfv136
7. Yahalom G, Schwartz R, Schwammenthal Y *et al*. Chronic kidney disease and clinical outcome in patients with acute stroke. *Stroke* 2009;**40**:1296–303. https://doi.org/10.1161/STROKEAHA.108.520882
8. Gage BF, Waterman AD, Shannon W *et al*. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. *JAMA* 2001;**285**:2864–70. https://doi.org/10.1001/jama.285.22.2864
9. Lip GY, Nieuwlaat R, Pisters R *et al*. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest* 2010;**137**:263–72. https://doi.org/10.1378/chest.09-1584
10. Olesen JB, Lip GY, Hansen ML *et al*. Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. *BMJ* 2011;**342**:d124. https://doi.org/10.1136/bmj.d124
11. Fang MC, Singer DE, Chang Y *et al*. Gender differences in the risk of ischemic stroke and peripheral embolism in atrial fibrillation: the AnTicoagulation and Risk factors In Atrial fibrillation (ATRIA) study. *Circulation* 2005;**112**:1687–91. https://doi.org/10.1161/CIRCULATIONAHA.105.553438
12. Pisters R, Lane DA, Nieuwlaat R *et al*. 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**:1093–100. https://doi.org/10.1378/chest.10-0134
13. 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**:713–9. https://doi.org/10.1016/j.ahj.2005.04.017
14. Page MJ, McKenzie JE, Bossuyt PM *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev* 2021;**10**:89. https://doi.org/10.1186/s13643-021-01626-4
15. Moons KGM, de Groot JAH, Bouwmeester W *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014;**11**:e1001744. https://doi.org/10.1371/journal.pmed.1001744
16. Wolff RF, Moons KGM, Riley RD *et al*. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;**170**:51–58. https://doi.org/10.7326/M18-1376
17. Guyatt GH, Oxman AD, Vist GE *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**:924–6. https://doi.org/10.1136/bmj.39489.470347.AD
18. Debray TP, Damen JA, Snell KI *et al*. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;**356**:i6460. https://doi.org/10.1136/bmj.i6460
19. Wan X, Wang W, Liu J *et al*. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014;**14**:135. https://doi.org/10.1186/1471-2288-14-135
20. Nadarajah R, Younsi T, Romer E *et al*. Prediction models for heart failure in the community: a systematic review and meta-analysis. *Eur J Heart Fail* 2023;**25**:1724–38. https://doi.org/10.1002/ejhf.2970
21. Debray TP, Damen JA, Riley RD *et al*. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. *Stat Methods Med Res* 2019;**28**:2768–86. https://doi.org/10.1177/0962280218785504
22. Himmelreich JCL, Veelers L, Lucassen WAM *et al*. Prediction models for atrial fibrillation applicable in the community: a systematic review and meta-analysis. *EP Europace* 2020;**22**:684–94. https://doi.org/10.1093/europace/euaa005
23. He Y, Ong Y, Li X *et al*. Performance of prediction models on survival outcomes of colorectal cancer with surgical resection: a systematic review and meta-analysis. *Surg Oncol* 2019;**29**:196–202. https://doi.org/10.1016/j.suronc.2019.05.014
24. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation* 2010;**121**:1768–77. https://doi.org/10.1161/CIRCULATIONAHA.109.849166
25. Nopp S, Spielvogel CP, Schmaldienst S *et al*. Bleeding risk assessment in end-stage kidney disease: validation of existing risk scores and evaluation of a machine learning-based approach. *Thromb Haemost* 2022;**122**:1558–66.
26. Ocak G, Ramspek C, Rookmaaker MB *et al*. Performance of bleeding risk scores in dialysis patients. *Nephrol Dial Transplant* 2019;**34**:1223–31. https://doi.org/10.1093/ndt/gfy387
27. Chao TF, Liu CJ, Wang KL *et al*. Incidence and prediction of ischemic stroke among atrial fibrillation patients with end-stage renal disease requiring dialysis. *Heart Rhythm* 2014;**11**:1752–9. https://doi.org/10.1016/j.hrthm.2014.06.021
28. Shih CJ, Ou SM, Chao PW *et al*. Risks of death and stroke in patients undergoing hemodialysis with new-onset atrial fibrillation: a competing-risk analysis of a nationwide cohort. *Circulation* 2016;**133**:265–72. https://doi.org/10.1161/CIRCULATIONAHA.115.018294
29. Sab M, Chelala D, Aoun M *et al*. Stroke in hemodialysis patients and its association with CHA(2)DS(2)-VASC and HAS-BLED scores: a retrospective study. *Clin Kidney J* 2023;**16**:596–602. https://doi.org/10.1093/ckj/sfac260
30. Wang TK, Sathananthan J, Marshall M *et al*. Relationships between anticoagulation, risk scores and adverse outcomes in dialysis patients with atrial fibrillation. *Heart Lung Circ* 2016;**25**:243–9. https://doi.org/10.1016/j.hlc.2015.08.012
31. de Jong Y, Ramspek CL, van der Endt VHW *et al*. A systematic review and external validation of stroke prediction models demonstrates poor performance in dialysis patients. *J Clin Epidemiol* 2020;**123**:69–79. https://doi.org/10.1016/j.jclinepi.2020.03.015

32. Madken M, Mallick R, Rhodes E *et al*. Development and validation of a predictive risk algorithm for bleeding in individuals on long-term hemodialysis: an international prospective cohort study (BLEED-HD). *Can J Kidney Health Dis* 2023;**10**:20543581231169610. https://doi.org/10.1177/20543581231169610
33. 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**:600–8. https://doi.org/10.1038/ki.2013.170
34. 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**:e006391. https://doi.org/10.1161/JAHA.117.006391.
35. 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**:91–9. https://doi.org/10.1016/S0002-9343(98)00198-3
36. Wang TJ, Massaro JM, Levy D *et al*. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the Framingham Heart Study. *JAMA* 2003;**290**:1049–56. https://doi.org/10.1001/jama.290.8.1049
37. Seliger SL, Gillen DL, Longstreth WT, Jr *et al*. Elevated risk of stroke among patients with end-stage renal disease. *Kidney Int* 2003;**64**:603–9. https://doi.org/10.1046/j.1523-1755.2003.00101.x
38. Masson P, Kelly PJ, Craig JC *et al*. Risk of stroke in patients with ESRD. *Clin J Am Soc Nephrol* 2015;**10**:1585–92. https://doi.org/10.2215/CJN.12001214
39. Molnar AO, Bota SE, Garg AX *et al*. The risk of major hemorrhage with CKD. *J Am Soc Nephrol* 2016;**27**:2825–32. https://doi.org/10.1681/ASN.2015050535
40. Zamberg I, Assouline-Reinmann M, Carrera E *et al*. Epidemiology, thrombolytic management, and outcomes of acute stroke among patients with chronic kidney disease: a systematic review and meta-analysis. *Nephrol Dial Transplant* 2022;**37**:1289–301. https://doi.org/10.1093/ndt/gfab197
41. Law MM, Tan SJ, Wong MCG *et al*. Atrial fibrillation in kidney failure: challenges in risk assessment and anticoagulation management. *Kidney Med* 2023;**5**:100690. https://doi.org/10.1016/j.xkme.2023.100690
42. Wizemann V, Tong L, Satayathum S *et al*. Atrial fibrillation in hemodialysis patients: clinical features and associations with anticoagulant therapy. *Kidney Int* 2010;**77**:1098–106. https://doi.org/10.1038/ki.2009.477
43. Dzeshka MS, Lane DA, Lip GY. Stroke and bleeding risk in atrial fibrillation: navigating the alphabet soup of risk-score acronyms ($CHADS_2$, CHA_2DS_2-VASc, R_2CHADS_2, HAS-BLED, ATRIA, and more). *Clin Cardiol* 2014;**37**:634–44. https://doi.org/10.1002/clc.22294
44. De Vriese AS, Heine G. Anticoagulation management in haemodialysis patients with atrial fibrillation: evidence and opinion. *Nephrol Dial Transplant* 2022;**37**:2072–9. https://doi.org/10.1093/ndt/gfab060
45. Doundoulakis I, Nedios S, Zafeiropoulos S *et al*. Atrial fibrillation burden: stepping beyond the categorical characterization. *Heart Rhythm* 2024:S1547–5271(24)03269-7. https://doi.org/10.1016/j.hrthm.2024.08.051
46. Kaplan RM, Koehler J, Ziegler PD *et al*. Stroke risk as a function of atrial fibrillation duration and CHA(2)DS(2)-VASc score. *Circulation* 2019;**140**:1639–46. https://doi.org/10.1161/CIRCULATIONAHA.119.041303
47. Boriani G, Botto GL, Padeletti L *et al*. Improving stroke risk stratification using the CHADS2 and CHA2DS2-VASc risk scores in patients with paroxysmal atrial fibrillation by continuous arrhythmia burden monitoring. *Stroke* 2011;**42**:1768–70. https://doi.org/10.1161/STROKEAHA.110.609297
48. Tiver KD, Quah J, Lahiri A *et al*. Atrial fibrillation burden: an update-the need for a CHA2DS2-VASc-AFBurden score. *Europace* 2021;**23**:665–73. https://doi.org/10.1093/europace/euaa287

Received: 17.10.2024; Editorial decision: 12.11.2024