Optimising prediction of mortality, stroke, and major bleeding for patients with atrial fibrillation

Apenteng, Patricia; Prieto-Merino , David; Hee , Siew Wan; Lobban , Trudie; Caleyachetty, Rishi; Fitzmaurice, David

DOI:
10.3399/BJGP.2023.0082

Document Version
Peer reviewed version

Citation for published version (Harvard):
Apenteng, P, Prieto-Merino , D, Hee , SW, Lobban , T, Caleyachetty, R & Fitzmaurice, D 2023, 'Optimising prediction of mortality, stroke, and major bleeding for patients with atrial fibrillation: external validation of the GARFIELD-AF model in UK primary care electronic records', *British Journal of General Practice* . https://doi.org/10.3399/BJGP.2023.0082

Accepted Manuscript

British Journal of General Practice

Optimising prediction of mortality, stroke, and major bleeding in patients with atrial fibrillation: external validation of the GARFIELD-AF model in UK primary care electronic records

Apenteng, Patricia; Prieto Merino, David; Hee, Siew Wan; Lobban, Trudie; Caleyachetty, Rishi; Fitzmaurice, David

DOI: https://doi.org/10.3399/BJGP.2023.0082

To access the most recent version of this article, please click the DOI URL in the line above.

Received 13 February 2023
Revised 26 May 2023
Accepted 15 June 2023

Optimising prediction of mortality, stroke, and major bleeding for patients with atrial fibrillation: external validation of the GARFIELD-AF model in UK primary care electronic records

Patricia N Apenteng,[1,2] Research Fellow, BA(Hons), MPhil, PhD

David Prieto-Merino,[3] Associate Professor, MSc, PhD

Siew Wan Hee,[2] Senior Research Fellow, PhD

Trudie Lobban,[4] Founder and Director, MBE FRCP

Rishi Caleyachetty,[5] GP ST1 trainee, MBBS, PhD

David A Fitzmaurice,[2] Professor of Cardiorespiratory Primary Care (Retired), MBChB, MRCGP, MD, FRCGP

1. Institute of Applied Health Research, University of Birmingham, Birmingham, UK
2. Warwick Medical School, University of Warwick, Coventry, UK
3. Faculty of Medicine, University of Alcala, Madrid, Spain
4. Atrial Fibrillation Association, UK
5. Coventry and Warwickshire VTS, UK

Contact information for corresponding author:

Name: Dr Patricia Apenteng

Address: Institute of Applied Health Research, University of Birmingham, Edgbaston B15 2TT

Email: p.n.k.apenteng@bham.ac.uk

ORCID: https://orcid.org/0000-0003-0835-3495

Abstract

Background

The GARFIELD-AF tool is a novel risk tool that simultaneously assesses the risk of all-cause mortality, stroke or systemic embolism, and major bleeding in patients with atrial fibrillation (AF).

Aim

To validate the GARFIELD-AF tool in UK primary care electronic records

Design and setting

Retrospective cohort study using Clinical Practice Research Datalink (CPRD) linked with Hospital Episode Statistics data and Office for National Statistics mortality data.

Method

Discrimination was evaluated using the area under the curve (AUC) and calibration was evaluated using calibration-in-the-large regression and calibration plots.

Results

486,818 patients aged ≥18 years with incident diagnosis of non-valvular AF between 2 January 1998 and 31 July 2020 were included; 50.6% received anticoagulation at diagnosis. The GARFIELD-AF models outperformed the CHA_2DS_2VASc and HAS-BLED tools in discrimination ability of death, stroke, and major bleeding at all the time points. The AUC (95%CI) for events at 1 year for the 2017 model were: death 0.747 (0.744 to 0.751) vs 0.635

(0.631 to 0.639) for CHA_2DS_2VASc; stroke 0.666 (0.663 to 0.669) vs 0.625 (0.622 to 0.628) for CHA_2DS_2VASc; and major bleeding 0.602 (0.598 to 0.606) vs 0.558; (0.554 to 0.562) **for HAS-BLED**. Calibration between predicted and Kaplan-Meier observed events was inadequate.

Conclusions

The GARFIELD models were superior to the CHA_2DS_2VASc score for discriminating stroke and death and to the HAS-BLED score for discriminating major bleeding. The models consistently under-predicted the level of risk, suggesting that a recalibration is needed to optimise its use in the UK population.

Keywords: atrial fibrillation, stroke, anticoagulation, all-cause mortality, bleeding, risk stratification

Word count: 2660

What is already known on this topic

- Anticoagulation reduces the risk of AF-related stroke at the cost of an increased risk of bleeding
- The CHA_2DS_2VASc tool is used to assess stroke risk in patients with AF, whilst the either the HAS-BLED or ORBIT-AF tool is used to assess bleeding risk
- A novel tool, GARFIELD-AF simultaneously predicts the risk of stroke death and bleeding in patients with AF, however its performance has not been tested in the UK population

What this study adds

- The GARFIELD-AF tool had better discriminatory ability than the CHA_2DS_2VASc and HAS-BLED in the UK population, however it underestimated the level of risk

Introduction

Oral anticoagulation (OAC) substantially reduces the risk of AF-related stroke.[1] However, OAC increases the risk of bleeding, and AF management guidelines recommend the use of risk stratification tools to guide decisions on anticoagulation.[2 3] The European Society of Cardiology (ESC) and the National Institute for Health and Care Excellence (NICE) AF guidelines recommend the CHA_2DS_2VASc score for assessing stroke risk.[3] Until recently both guidelines recommended the HAS-BLED tool for accessing bleeding risk; however since 2021 NICE recommends the ORBIT-AF risk score.[3]

The recommended tools are widely used in clinical practice, nevertheless up to 15% of patients with AF at risk of stroke in England do not receive guideline recommended therapy.[4] The GARFIELD-AF risk tool is a novel risk tool that simultaneously assesses a patient's risk of mortality, stroke or systemic embolism, and risk of major bleeding.[5 6] The GARFIELD-AF tool was developed based on 39,898 patients enrolled on the GARFIELD-AF registry in 2017[5] and a new version published in 2021 predicted events up to 2 years from diagnosis.[6] Initial evaluations indicate that both versions are superior to CHA_2DS_2VASc in predicting ischemic stroke/systemic embolism and HAS-BLED in predicting bleeding risk.[5 6]

GARFIELD-AF is an international prospective observational study of patients≥ 18 years with newly diagnosed AF and ≥1 investigator determined risk factor for stroke.[7 8] 52,080 participants were enrolled in in 35 countries and followed for a minimum of two years; 3574 of the GARFIELD-AF cohort were recruited in the UK.[9] The GARFIELD-AF tool can potentially be embedded into primary care electronic systems to aid decision-making regarding anticoagulation so that patients who require anticoagulation receive it and those that do not need it do not receive it.

The performance of prediction model tends to vary across settings and populations, and external validation is required to fully appreciate the generalisability of a prediction model.[10 11] The purpose of this study was to validate the GARFIELD-AF tool in patients with AF in an NHS primary care electronic health records database, and compare its performance of the GARFIELD-AF tool with the CHA_2DS_2VASc and HAS-BLED tools.

Methods

Source of data

The primary data source was the Clinical Practice Research Datalink (CPRD) an electronic primary care database comprising anonymised patient medical records from general practitioners, with coverage of over 19 million patients from 738 practices in the UK.[12] Data were extracted by CPRD and linked with Hospital Episode Statistics (HES) data which provides information on all hospital admissions and mortality data from Office for National Statistics (ONS).

Study population

The study population was defined as adults aged ≥18 years, with incident diagnosis of non-valvular AF between 2 January 1998 and 31 July 2020, and eligible for linkage with HES and ONS data.

Follow up

Start of follow up was defined as the recorded date the patient was diagnosed with non-valvular AF. End of follow up was defined as death as recorded by ONS, end of practice registration or last collection date, whichever occurred first.

Covariates

The covariates for the GARFIELD-AF models are: age, sex, pulse, systolic blood pressure (SBP) and diastolic blood pressure (DBP), weight, height, ethnicity, current smoking, paroxysmal AF; history of vascular disease, diabetes, cirrhosis, peripheral vascular disease, stroke, bleeding, heart failure, chronic kidney disease, sleep apnoea, dementia, carotid occlusive disease; and anticoagulant use and antiplatelet use. The covariates and coefficients for the 2017 and 2021 models are detailed in Supplementary Tables S1 and S2. The main difference between the 2017 and the 2021 models is that the 2021 models have a wider range of variables. For example, the 2017 GARFIELD-AF model for stroke includes the variables age, history of stroke, bleeding, heart failure, chronic kidney disease, regional and ethnicity, and anticoagulant use. The 2021 GARFIELD-AF model for stroke has the additional variables female sex, history of carotid exclusive disease, dementia, and smoking. The GARFIELD-AF 2017 models for death has a full version and a simpler version that comprises of a reduced set of variables (age, pulse, SBP, a history of vascular disease, history of bleeding, heart failure, renal disease, and anticoagulant use), whereas the 2021 death model has just one version.

The covariates for the CHA_2DS_2VASc score are: history of congestive heart failure, hypertension, age, diabetes, prior stroke, vascular disease, and sex. The covariates for the HAS-BLED score are: hypertension, abnormal liver or renal function, history of stroke,

bleeding history, labile INR, age, drugs use at time of diagnosis (antiplatelets or NSAIDs) or alcohol use.

The baseline variables for the GARFIELD-AF models and CHA_2DS_2VASc and HAS-BLED were defined from CPRD data using Medical Code IDs. Details are provided in Supplementary Box 1.

Definition of endpoints

The study endpoints were all-cause mortality; ischaemic stroke/systemic embolism (SE), defined as the combined end point of any ischaemic stroke, transient ischaemic attack (TIA), or SE; and major bleeding (including haemorrhagic stroke), defined as bleeding requiring hospitalisation. The first occurrence of an ischaemic stroke/SE after AF diagnosis was the endpoint for ischaemic stroke/SE and the first occurrence of a major bleeding after AF diagnosis was the endpoint for major bleeding.

Outcome variables

Outcome variables were defined from both Medical code IDs and ICD-10 codes for HES and ONS mortality data as detailed in Supplementary Box 2.

Statistical analysis

The GARFIELD models were applied to the CPRD dataset to obtain the predicted risks for each outcome. The performance of the tool was measured in terms of calibration using calibration-in-the-large regression and calibration plots, and in terms of discrimination using the area under the receiver operating characteristic curve (AUC), also referred to as the C-

statistic. The performance of the models was compared with the CHA_2DS_2VASc and HAS-BLED tools by comparing AUC of each model. The CHA_2DS_2VASc tool, in addition to predicting the risk of stroke in patients with AF, has been shown to predict mortality in patients with several diseases, regardless of the presence of AF.[13] The performance of the CHA_2DS_2VASc tool for predicting stroke and death was compared to the GARFELD-AF models for stroke and death and the performance of HAS-BLED for predicting bleeding was compared to the GARFIELD-AF bleeding models. The treatment effect was estimated by running separate Cox regression models for each outcome (death, stroke, and bleeding) and adjusting each model for all the variables that contribute to the GARFIELD 2021 score for that outcome.

Each variable was assessed for the degree of missingness. The assessment for discrimination and calibration was performed on the whole dataset and repeated in patients without missing data in any score. Subgroup analysis was conducted according to risk stratification of stroke (high, moderate, and low according to CHA_2DS_2VASc) and bleeding (HAS-BLED less than or more than 2) and for individuals receiving anticoagulation or no anticoagulation at baseline.

Results

A total of 708,474 patients had an incident record of AF in CPRD Aurum. Of these, 486,818 met the inclusion criteria for the study (Figure 1). The median follow-up was 3.975 years (IQR, 1.6 to 7.7 years; min 0 year to max 22.6 years).

Baseline characteristic of participants

The baseline characteristics for the CPRD validation cohort, the UK GARFIELD-AF sub-cohort and the global GARFIELD-AF cohort are presented in Table 1. The UK cohorts were older compared to the global GARFIELD-AF cohort (mean age of 75 years vs 70 years). The UK cohorts were predominantly of white ethnicity (95% CPRD and 99% GARFIELD UK vs 63% global cohort) and had a higher prevalence of history of bleeding (6.8% CPRD vs 2.5% global cohort).

Four-fifths of the CPRD cohort had $CHA_2DS_2VASc \geq 2$, 8.5% had $CHA_2DS_2VASc=1$ and 8.2% had $CHA_2DS_2VASc=0$. 50.6% of the CPRD cohort received anticoagulation at diagnosis compared to 65.8% in the UK GARFIELD cohort and 66.9% in the global GARFIELD cohort. Overall, the CPRD cohort had a lower mean CHA_2DS_2VASc score compared to the GARFIELD UK and global cohorts: 2.96 (standard deviation, SD, 1.5) vs 3.3 (SD, 1.5) and 3.2 (SD, 1.6), respectively. In the CPRD cohort 83.02% of the patients with a HAS-BLED score<3 compared to 81.07% in GARFIELD UK and 88.77% in the global GARFIELD cohort.

Missing data in CPRD

There were no missing data in the covariates needed to calculate CHA_2DS_2VASc, but 33% (162,298/486,818) patients had missing data for calculating HAS-BLED. For the 2017 GARFIELD-AF models there were no missing data for the predictors for the bleeding model, but 16,075 patients (3%) had missing data for the stroke model and 67% had missing data for the mortality model. Therefore, we could calculate all three models (bleeding, stroke, and mortality) in only 164,427 patients (34%). For the 2021 GARFIELD-AF models, 69% of patients had missing data for the bleeding model, 65% had missing data for the stroke

model, and 89% had missing data for the mortality model. 53,228 patients (11%) had complete data for all three models (bleeding, stroke, and mortality).

External Validation for the GARFIELD-AF models

Table 2 shows the full data for the 2017 1-year mortality, stroke and bleeding models and the 2021 models (each model with 1-month, 1-year and 2-year follow-up).

Discrimination

The AUCs in table-2 range from 0.576 [95% CI, 0.565 to 0.586] of the 2021 model for bleeding in 1 month to 0.753 [95% CI, 0.737 to 0.769] of the 2021 model for death in 1 month. At 1-year follow-up the 2017 and 2021 models performed very similar for the outcomes of stroke (AUCs 0.666 vs 0.670, respectively) and bleeding (AUCs 0.603 vs 0.598, respectively) with overlapping confidence intervals in both outcomes. The 1-year 2017 model slightly but significantly outperformed the 1-year 2021 model for predicting death with AUCs 0.747 [95% CI, 0.744 to 0.751] and 0.728 [95% CI, 0.722 to 0.735], respectively, with non-overlapping confidence intervals.

Calibration

In the three outcomes, both the 2017 and 2021 GARFIELD AF models consistently predicted less average risk than the observed risks in the population estimated using the Kaplan-Meier (KM) method (table 2 and Figure 2). The 2017 model performed slightly better than the 2021 model at 1-year follow-up in the three outcomes. The calibration plots shows that the differences between the GARFIELD's predicted risks and the KM estimated risks grow in the larger quintiles (Figure 3).

Comparison GARFIELD models and CHA_2DS_2VASc and HAS-BLED scores

The GARFIELD models consistently outperformed the CHA_2DS_2VASc and HAS-BLED scores (Table 3). The AUC for the 2017 models at 1-year follow-up were: death 0.748 vs 0.635 for CHA_2DS_2VASc, stroke 0.666 vs 0.625 for CHA_2DS_2VASc and bleeding 0.602 vs 0.558 for HAS-BLED. The AUC for the 2021 models at 1 year were death: 0.728 vs 0.616 for CHA_2DS_2VASc, stroke 0.670 vs 0.620 for CHA_2DS_2VASc and bleeding 0.604 vs 0.560 for HAS-BLED. P-values were less than <0.00001 with non-overlapping confidence intervals in all comparisons.

Subgroup analyses

Patients not taking OAC showed a higher average risk of events in almost every version of the model than those patients taking OAC. The AUC was always larger in patients not taking OAC. This is compatible with OAC lowering the risks of patients and making it more difficult to tell who is going to have an event (lower AUC) (Table S3).

After adjusting for the GARFIELD 2021 risk factors in Cox regression models, anticoagulation had a protective effect from death with a hazard ratio of aHR=0.58 (95% CI, 0.50 to 0.68), a protective effect for stroke aHR=0.71 (95% CI, 0.63 to 0.81) and a non-significant protective effect on bleeding aHR= 0.90 (95% CI, 0.76 to 1.05).

When stratified according to risk levels, the GARFIELD tools performed better in patients at high risk compared to moderate risk for stroke according to CHA_2DS_2VASc (Table S3). The AUCs for 2017 1-year risk for stroke model were: high risk 0.652 (95% CI, 0.649 to 0.656), moderate risk 0.559 (95% CI, 0.545 to 0.572), and low risk 0.526 (95% CI, 0.508 to 0.543).

Complete case analysis

The data for the complete case analysis is shown in Table S4. When analysis was restricted to the 30,666 patients with data to calculate all scores the AUCs for the 2017 model were: death 0.719 (95% CI, 0.710 to 0.728), stroke 0.677 (95% CI, 0.665 to 0.689) and bleed was 0.589 (95% CI, 0.573 to 0.598), indicating a similar performance to the whole group analysis for stroke and bleed and a slight difference for death. Like in the main analyses, the models were miscalibrated when restricted to cases with patients with full dataset, showing important differences between the GARFIELD-predicted and Kaplan-Meier estimated risks (Figure S1).

Discussion

Summary

In this study population of 486,818 patients with incident AF, the GARFIELD models have good discrimination for predicting death and moderate discrimination for predicting stroke and bleeding, but consistently below the discriminations reported in the original GARFIELD publications. Our findings show that the models are superior to the CHA_2DS_2VASc score for predicting stroke and the HAS-BLED score for predicting bleeding. However, all versions of the model consistently under-predicted the level of risk. There were no significant differences in the performance of the 2017 and 2021 models at 1 year for bleeding and stroke but model 2017 showed a slightly better performance for death.

Strengths and limitations

The study has several strengths, the data source CPRD is a primary care database representing approximately 10% of the UK primary care population, provided a large real-world sample of patients with incident AF with good statistical power. Linkage with HES and ONS data improved robustness of the data, reducing chance of missing out any of the outcomes of interest. We used a rigorous process of codelist development, with a primary care clinician overseeing and reviewing all the codelists.

There was a significant amount of missing data for the 2021 models and the 2017 full death model. The volume of missing data was too large for multiple imputation, comparison of whole dataset and complete cases shown little difference.

Comparison with existing literature

Within the global GARFIELD-AF study population the 2017 model had a modest predictive ability for stroke (0.69 [95% CI, 0.67 to 0.71]) and major bleeding (0.66 [95% CI, 0.62 to 0.69]) and a good performance for death (0.77 [95% CI, 0.76 to 0.78]).[5] These values were slighter better when the 2021 models were evaluated in the GARFIELD-AF study population, the AUC at 1 year were: stroke 0.70. (95% CI, 0.68 to 0.72), major bleeding 0.69 (95% CI, 0.67 to 0.71) and death 0.76 (95% CI, 0.75 to 0.77).[6] In both the 2017 and 2021 internal validation the models performed better than within the UK CPRD cohort; however this is what one would expect of an internal validation.

External validation of the GARFIELD-AF model

The GARFIELD-AF models performed better in the ORBIT-AF population than in the UK; AUC death 0.75 (95% CI, 0.74 to 0.76); stroke 0.68 (95% CI, 0.64 to 0.71) major bleeding 0.64 (95% CI, 0.62 to 0.66) for 2021 model[6] and similar outcomes for the 2017 models[5].

Our study is the first to independently validate the GARFIELD-AF 2021 models externally, however the 2017 models have been previously evaluated. An independent evaluation of the GARFIELD-AF 2017 stroke and bleeding in the Danish population reported higher discriminatory than we found in the CPRD cohort: stroke (AUC 0.71 (95% CI, 0.70 to 0.72); HAS-BLED 0.64 (95% CI, 0.63 to 0.66) in patients using OAC therapy.[14] The same study found that the GARFIELD-AF model was superior to CHA_2DS_2VASC but comparable to HAS-BLED whereas in our study population discrimination ability of GARFIELD was superior to both CHA_2DS_2VASC and HAS-BLED.[14] In addition both the stroke and bleeding models were well calibrated in the Danish cohort.[14] Another study comparing the 2017 GARFIELD bleeding score to HAS-BLED also found bleeding model to have modest predictive value though the c- statistic for GARFIELD-AF was lower than in our study 0.56 (95% CI, 0.54 to 0.57).[15]

The GARFIELD-AF models consistently under-predicted the level of risk in the CPRD cohort. There may be a number of reasons for this – it may be the impact of geographical variation; there were significant variations in outcomes across countries within the GARFIELD-AF registry even after adjustment for baseline characteristics and antithrombotic treatment.[16] There were differences in the baseline characteristics in the UK population, notably the UK population were older. Also the UK GARFIELD population a higher a incidence of stroke, bleeding and mortality compared to the global population.[9]

Implications for research and practice

The novelty of the GARFIELD-AF model is simultaneous prediction of stroke, bleeding, and death. Death has been shown to be an important outcome in atrial fibrillation, prompting recommendations for a more integrated management of patients with AF,[2] however there is currently no designated tool for assessing mortality in patients with AF. The death models had the best predictive ability, and the 2017 abridged death model offers a good alternative with a reduced set of predictors that are available in UK primary are records.

The 2017 model would be better suited to clinical use in the UK due to better availability of the predictors in primary care records. Recalibration will optimise the use of the GARFIELD-AF model in the UK population without losing the information captured from the original model. Incorporating a recalibrated tool into UK primary care electronic systems would help clinicians evaluate the risk benefit ratio of anticoagulation and potentially improve risk stratification and decision-making regarding anticoagulation in patients with AF.

Funding

This project is funded by the National Institute for Health and Care Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (Grant Reference Number NIHR200831). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Ethical approval

Ethic approval was obtained from the Warwick Medical School Biomedical & Scientific Research Ethics Committee (BSREC) - Ref BSREC 26/19-20.

Access to CPRD data is subject to protocol approval by an Independent Scientific Advisory Committee, a non-statutory expert advisory body oversees access to linked CPRD data for research purposes. Approval was obtained from the CPRD Independent Scientific Advisory Committee (Ref 19_276).

Competing interests

The authors report no competing interests.

Acknowledgements

We thank our collaborators Saverio Virdone (Thrombosis Research Institute, London), Karen Pieper (Thrombosis Research Institute, London) and Professor Keith Fox (University of Edinburgh) for methodological input and feedback throughout the study.

PA is affiliated to the National Institute for Health Research (NIHR) Applied Research Collaboration (ARC) West Midlands. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Open Access

References

1. The Atrial Fibrillation Investigators. Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation: analysis of pooted data from five randominized controlled trials: *Arch Intern Med* 1994;154(13):1449-14.
2. 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* 2020
3. National Institute for Health and Care Excellence Atrial fibrillation: diagnosis and management NICE guideline, 2021.
4. NHS Digital. Quality and Outcomes Framework, Achievement, prevalence and exceptions data 2018-19 Available from https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-prevalence-and-exceptions-data/2018-19-pas, 2019.
5. Fox KA, Lucas JE, Pieper KS, et al. Improved risk stratification of patients with atrial fibrillation: an integrated GARFIELD-AF tool for the prediction of mortality, stroke and bleed in patients with and without anticoagulation. *BMJ Open* 2017;7(12):e017157.
6. Fox KA, Virdone S, Pieper KS, et al. GARFIELD-AF risk score for mortality, stroke and bleeding within 2 years in patients with atrial fibrillation. *Eur Heart J Qual Care Clin Outcomes* 2022;8(2):214-27.
7. Kakkar AK, Mueller I, Bassand JP, et al. International longitudinal registry of patients with atrial fibrillation at risk of stroke: Global Anticoagulant Registry in the FIELD (GARFIELD). *Am Heart J* 2012;163(1):13-19.
8. Apenteng PN, Murray ET, Holder R, et al. An international longitudinal registry of patients with atrial fibrillation at risk of stroke (GARFIELD): the UK protocol. *BMC Cardiovasc Disord* 2013;13(1):31. doi: 10.1186/1471-2261-13-31
9. Apenteng P, Virdone S, Hobbs FR, et al. Two-year outcomes of UK patients newly diagnosed with atrial fibrillation: findings from the prospective observational cohort study GARFIELD-AF. *Br J Gen Pract* 2022:BJGP.2021.0548. doi: 10.3399/bjgp.2021.0548
10. Bleeker S, Moll H, Steyerberg E, et al. External validation is necessary in prediction research: A clinical example. *J Clin Epidemiol* 2003;56(9):826-32.
11. Riley RD, Ensor J, Snell KI, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 2016;353:i3140.
12. Wolf A, Dedman D, Campbell J, et al. Data resource profile: clinical practice research Datalink (CPRD) aurum. *Int J Epidemiol* 2019;48(6):1740-40g.
13. Goudis C, Daios S, Korantzopoulos P, Liu T. Does CHA2DS2-VASc score predict mortality in chronic kidney disease? *Intern Emerg Med.* 2021 Oct;16(7):1737-1742. doi: 10.1007/s11739-021-02799-5. Epub 2021 Jul 7. PMID: 34232486; PMCID: PMC8261034.
14. Dalgaard F, Pieper K, Verheugt F, et al. GARFIELD-AF model for prediction of stroke and major bleeding in atrial fibrillation: a Danish nationwide validation study. *BMJ Open* 2019;9(11):e033283.
15. Proietti M, Rivera-Caravaca JM, Esteve-Pastor MA, et al. Predicting bleeding events in anticoagulated patients with atrial fibrillation: a comparison between the HAS-BLED and GARFIELD-AF bleeding scores. *J Am Heart Assoc* 2018;7(18):e009766.
16. Fox KA, Virdone S, Bassand J-P, et al. Do baseline characteristics and treatments account for geographical disparities in the outcomes of patients with newly diagnosed atrial fibrillation? The prospective GARFIELD-AF registry. *BMJ Open* 2022;12(1):e049933.

Tables and Figures

Supplementary file

Table 1. Baseline characteristics of the CPRD validation cohort, the UK GARFIELD cohort and the global GARFIELD-AF cohort

Variable	Validation cohort	GARFIELD UK	Global GARFIELD
N	**486818**	**3574**	**52080**
Age, years, mean (SD)	74.6 (12.2)	74.5 (9.5)	69.7 (11.5)
Age <65, n (%)	89281 (18.3)	471 (13.2)	15708 (30.2)
Age 65 – 74, n (%)	123661 (25.4)	1178 (32.9)	16960 (32.6)
Age ≥75, n (%)	273876 (56.3)	1925 (53.9)	19412 (37.3)
Female, n (%)	227370 (46.7)	1522 (42.6)	23011 (44.2)
Ethnicity (N)	**470743**	**3483**	**50796**
White ethnicity, n (%)	447972 (95.2)	3441 (98.8)	32028 (63.1)
Asian ethnicity	7373 (1.6)	13 (0.4)	14302 (28.2)
Black/mixed/other ethnicity	15398 (3.3)	29 (0.8)	4466 (8.8)
Clinical observations at diagnosis			
Pulse, mean (SD)	79.2 (18.9)	87.5 (22.7)	90.4 (26.7)
Systolic blood pressure, mean (SD)	134.8 (19.6)	133.0 (17.7)	133.5 (19.8)
Diastolic blood pressure, mean (SD)	77.4 (11.7)	77.0 (11.3)	79.7 (12.9)
Weight (kg), mean (SD)	81.3 (21.0)	83.6 (19.7)	77.6 (19.0)
Height (m), mean (SD)	1.68 (0.1)	1.69 (0.10)	1.67 (0.10)
BMI, mean (SD)	28.8 (6.3)	29.2 (6.2)	27.8 (5.7)
Medical history, n (%)			
Congestive heart failure	33817 (6.95)	274 (7.7)	11758 (22.6)
History of hypertension	344590 (70.8)	2483 (69.7)	39643 (76.3)
Diabetes mellitus	66876 (13.7)	629 (17.6)	11555 (22.2)
Prior Stroke/TIA	52103 (10.7)	450 (12.7)	5961 (11.4)
Vascular disease	41523 (8.5)	760 (21.4)	7682 (14.8)
Peripheral vascular disease	10318 (2.1)	-	
Carotid occlusive disease	2107 (0.4)	52 (1.5)	1545 (3.0)
History of bleeding	33205 (6.8)	109 (3.1)	1318 (2.5)
Chronic kidney disease (grade ≥3)	69676 (14.3)	896 (25.6)	5360 (10.3)
Chronic renal failure	7831 (1.6)	-	-
Cirrhosis	1484 (0.3)	11 (0.3)	295 (0.6)
Current smoker	30168 (6.2)	245 (7.0)	5204 (11.0)
Sleep apnoea	5048 (1.0)	-	-
Dementia	10830 (2.2)	28 (0.8)	764 (1.5)
Type of AF diagnosed is paroxysmal	65474 (13.5)	651 (18.2)	14315 (27.5)
Antiplatelets or NSAIDs use	192271 (39.5)	-	-
≤8 units alcohol/week	362641 (74.5)		
>8 units alcohol/week	34513 (7.1)		
Risk scores			
CHA_2DS_2VASc score, mean (SD)	2.96 (1.5)	3.3 (1.5)	3.2 (1.6)
CHA_2DS_2VASc score categories, N	**486818**	**3528**	**51408**
0	40052 (8.2)	62 (1.8)	1516 (2.9)
1	41135 (8.5)	316 (9.0)	6369 (12.4)
2	88455 (18.2)	659 (18.7)	10230 (19.9)
3	126160 (25.9)	972 (27.6)	12138 (23.6)
4	124665 (25.6)	848 (24.0)	11022 (21.4)

5	51858 (10.7)	398 (11.3)	5895 (11.5)
≥6	14723 (2.93)	273 (7.7)	4238 (8.2)
HAS-BLED[1] score, mean (SD)	1.62 (0.9)	1.7 (0.9)	1.4 (0.9)
HAS-BLED[1] score categories, N	**324520**	**2530**	**37549**
0	30770 (6.3)	160 (6.3)	5471 (14.6)
1	125541 (25.8)	941 (37.2)	16169 (43.1)
2	113120 (23.2)	950 (37.5)	11692 (31.1)
3	45954 (9.4)	391 (15.5)	3570 (9.5)
≥4	9135 (1.9)	88 (3.5)	647 (1.7)
Treatment at diagnosis	**486818**	**3564**	**51354**
NOAC	106994 (22.0)	688 (19.3)	14129 (27.5)
VKA	141200 (29.0)	1656 (46.3)	20206 (39.3)
OAC	246425 (50.6)	2344 (65.8)	34335 (66.9)
AP	187962 (38.6)	1189 (33.4)	18121 (35.3)

[1]The risk factor 'Labile INRs' is not included in the HAS-BLED score. As a result, the maximum HAS-BLED score at baseline is 8 points (not 9); (2) Denominators of the medical history risk factors vary depending on how many individuals had the information available, the percentages are calculated on the number of people with information in each risk factor (not shown).

Table 2. Predicted and Kaplan-Meier estimated risks for the GARFIELD models

Model = Year of GARFIELD-AF model, Months = months of follow-up, Np = Number of patients with predicted risk, Pred = Average Predicted risk, KM = Estimated risk using Kaplan-Meier's method, N0 = Number of patients without the outcome at end of follow-up, P0 = Average predicted risk in patients in N0, N1 = Number of patients with positive outcome at end of follow-up, P1 = Average predicted risk in patients in N1, AUC = Area Under the Curve (C-statistic)

Model	Outcome	Months	Np	Pred	KM	N0	P0	N1	P1	AUC (95% CI)
2017	death	12	167197	4.51%	11.89%	130871	4.03%	19110	8.16%	0.748 [0.744 to 0.751]
2017(full)	death	12	107404	6.81%	10.81%	84574	6.09%	11132	12.76%	0.748 [0.743 to 0.752]
2017	stroke	12	470743	1.65%	8.29%	357601	1.48%	37466	2.32%	0.666 [0.663 to 0.669]
2017	bleed	12	486818	1.40%	6.32%	377482	1.31%	28521	1.61%	0.603 [0.599 to 0.606]
2021	death	1	53228	0.81%	1.55%	52218	0.79%	823	1.71%	0.753 [0.737 to 0.769]
2021	stroke	1	149105	0.22%	4.27%	140940	0.22%	6358	0.29%	0.622 [0.615 to 0.629]
2021	bleed	1	170123	0.22%	1.60%	164732	0.21%	2706	0.25%	0.576 [0.565 to 0.586]
2021	death	12	53228	5.98%	11.82%	41652	5.41%	6046	10.64%	0.728 [0.722 to 0.735]
2021	stroke	12	149105	1.58%	8.28%	111751	1.43%	11795	2.17%	0.670 [0.665 to 0.676]
2021	bleed	12	170123	1.55%	7.25%	123287	1.48%	11250	1.85%	0.598 [0.593 to 0.604]
2021	death	24	53228	10.27%	19.64%	32016	8.89%	9445	17.02%	0.731 [0.726 to 0.737]
2021	stroke	24	149105	2.62%	11.10%	89415	2.28%	14925	3.56%	0.683 [0.678 to 0.687]
2021	bleed	24	170123	2.41%	11.66%	91111	2.23%	16425	2.81%	0.602 [0.598 to 0.607]

Table 3. Comparison of GARFIELD models and CHA_2DS_2VASc (CHAD) and HAS-BLED (HASB)

Months = months of follow-up, AUC-GAR =AUC of the GARFIELD model in the row, AUC-other = AUC of the other model in the row, Pval = P-value comparing AUCs

GARFIELD	Other	Outcome	Months	AUC-GAR	AUC-other	Pval
2017	CHAD	death	12	0.748 [0.744 to 0.748]	0.635 [0.631 to 0.635]	<0.00001
2017(full)	CHAD	death	12	0.748 [0.743 to 0.748]	0.627 [0.622 to 0.627]	<0.00001
2017	CHAD	stroke	12	0.666 [0.663 to 0.666]	0.625 [0.622 to 0.625]	<0.00001
2017	HASB	bleed	12	0.602 [0.598 to 0.602]	0.558 [0.554 to 0.558]	<0.00001
2021	CHAD	death	1	0.753 [0.737 to 0.753]	0.609 [0.591 to 0.609]	<0.00001
2021	CHAD	stroke	1	0.622 [0.615 to 0.622]	0.588 [0.581 to 0.588]	<0.00001
2021	HASB	bleed	1	0.584 [0.571 to 0.584]	0.549 [0.538 to 0.549]	<0.00001
2021	CHAD	death	12	0.728 [0.722 to 0.728]	0.616 [0.609 to 0.616]	<0.00001
2021	CHAD	stroke	12	0.670 [0.665 to 0.670]	0.620 [0.615 to 0.620]	<0.00001
2021	HASB	bleed	12	0.604 [0.598 to 0.604]	0.560 [0.554 to 0.560]	<0.00001
2021	CHAD	death	24	0.731 [0.726 to 0.731]	0.625 [0.619 to 0.625]	<0.00001
2021	CHAD	stroke	24	0.683 [0.678 to 0.683]	0.634 [0.630 to 0.634]	<0.00001
2021	HASB	bleed	24	0.607 [0.602 to 0.607]	0.559 [0.554 to 0.559]	<0.00001

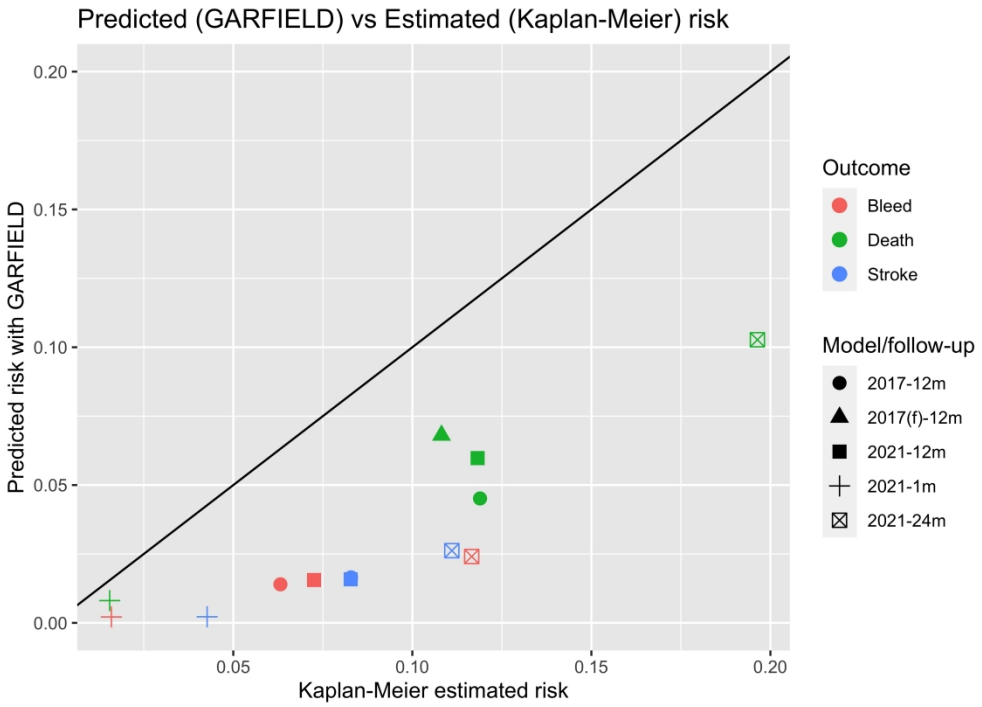

Figure 2. Predicted vs Kaplan-Meir estimated risks for the GARFIELD models

1481x1058mm (72 x 72 DPI)

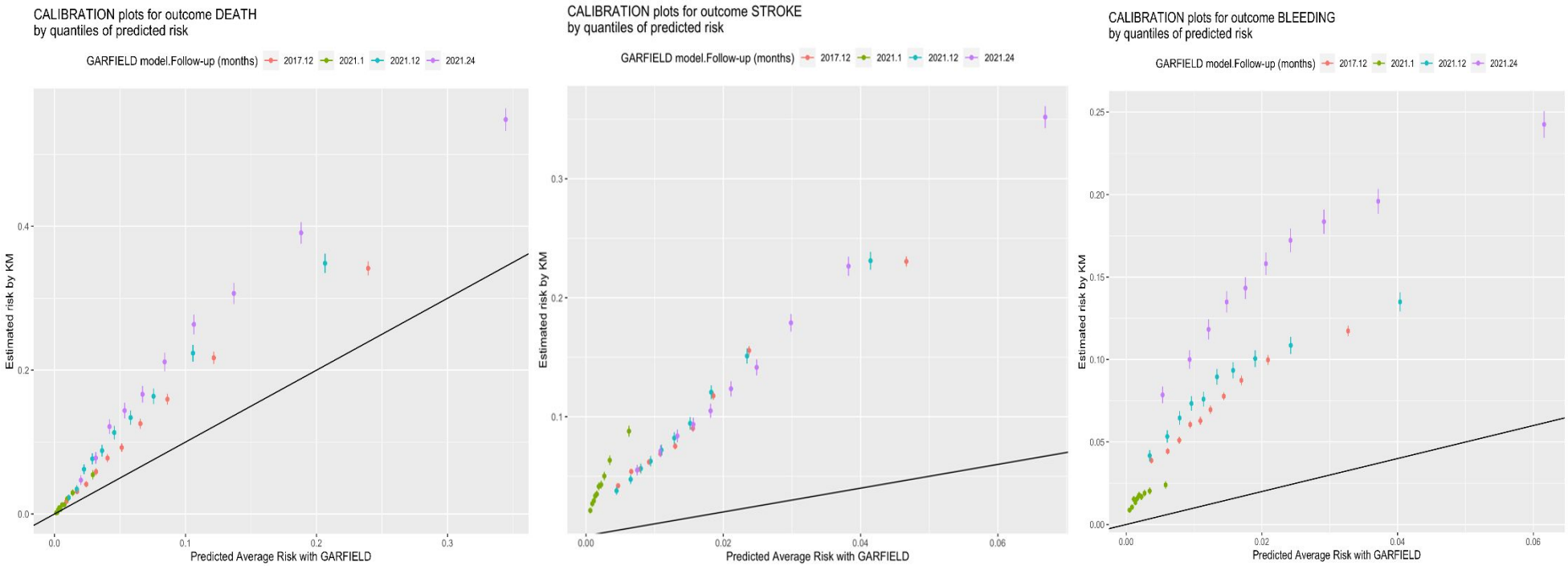

Figure 3. Calibration plots for death, stroke and bleeding outcomes